

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:31:01 ; Search time 109 Seconds

(without alignments)  
3131.118 Million cell updates/sec

Title: US-10-721-793-115

Perfect score: 192  
Sequence: 1 aaagacgggtatctggtgga.....ccctttctaataaagatgc 192

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq.\*
- 7: /cgn2\_6/prodata/1/ina/PP COMB.seq.\*
- 8: /cgn2\_6/prodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	121.4	63.2	301	3	US-09-053-021-3 Sequence 3, Appli
2	121.4	63.2	345	3	US-09-053-021-8 Sequence 8, Appli
3	35	18.2	243	3	US-09-599-632-11 Sequence 11, Appli
4	35	18.2	1664976	3	US-08-916-421B-1 Sequence 1, Appli
5	35	18.2	1664976	3	US-09-692-570-1 Sequence 1, Appli
6	32.4	16.9	3990	3	US-09-830-230A-629 Sequence 629, App
7	32.4	16.9	7766	3	US-09-125-619-3 Sequence 3, Appli
8	32.4	16.9	7766	3	US-10-222-566-3 Sequence 3, Appli
9	32.4	16.9	7766	3	US-10-143-024A-3 Sequence 3, Appli
10	32.4	16.9	7766	3	US-10-222-162-3 Sequence 17726, A
11	32	16.7	601	3	US-09-949-016-17726 Sequence 17726, A
12	32	16.7	601	3	US-09-949-016-17727 Sequence 17727, A
13	32	16.7	601	3	US-09-949-016-46726 Sequence 46726, A
14	32	16.7	601	3	US-09-949-016-46727 Sequence 46727, A
15	32	16.7	2560	3	US-09-023-655-1285 Sequence 1285, Ap
16	32	16.7	2578	3	US-09-949-016-1355 Sequence 1355, Ap
17	32	16.7	2591	3	US-09-949-016-9 Sequence 9, Appli
18	32	16.7	40742	3	US-09-949-016-11751 Sequence 11751, A
19	32	16.7	40747	3	US-09-949-016-13097 Sequence 13097, A
20	31.6	16.5	1059	3	US-09-419-788-26 Sequence 26, Appli
21	31.4	16.4	8286	3	US-09-949-016-15170 Sequence 15170, A
22	31.4	16.4	16082	3	US-09-949-016-16220 Sequence 16220, A
23	31.4	16.4	1664976	3	US-08-916-421B-1 Sequence 1, Appli
24	31.4	16.4	1664976	3	US-09-692-570-1 Sequence 1, Appli

25	30.2	15.7	1100	3	US-09-710-279-1201 Sequence 1201, Ap
26	30.2	15.7	1215	3	US-09-134-001C-277 Sequence 277, App
27	30.2	15.7	3153	3	US-09-710-279-3441 Sequence 3441, Ap
28	30	15.6	1329	3	US-09-248-796A-2713 Sequence 2713, Ap
29	29.6	15.4	996	3	US-09-134-001C-2351 Sequence 2351, Ap
30	29.6	15.4	4103	3	US-09-710-279-3532 Sequence 3532, Ap
31	29.4	15.3	194	3	US-09-445-803-5 Sequence 5, Appli
32	29.2	15.2	3394	3	US-09-710-279-3629 Sequence 3629, Ap
33	28.8	15.0	1293	3	US-09-248-796A-2701 Sequence 2701, Ap
34	28.8	15.0	119211	3	US-09-596-002-40 Sequence 40, Appli
35	28.6	14.9	531	3	US-09-543-681A-2750 Sequence 2750, Ap
36	28.6	14.9	3760	3	US-09-976-534-213 Sequence 213, App
37	28.6	14.9	49744	3	US-09-927-091-4 Sequence 4, Appli
38	28.4	14.8	267482	3	US-09-949-002-659 Sequence 659, App
39	28.4	14.8	267505	3	US-09-949-002-783 Sequence 783, App
40	28.2	14.7	250	3	US-09-513-999C-19914 Sequence 19914, A
41	28	14.6	643	3	US-09-034-088A-27 Sequence 27, Appli
42	28	14.6	643	3	US-09-781-811-27 Sequence 27, Appli
43	27.8	14.5	363	3	US-09-134-000C-1309 Sequence 1309, Ap
44	27.8	14.5	1960	2	US-08-595-937A-1 Sequence 1, Appli
45	27.8	14.5	2443	2	US-08-447-185-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-053-021-3  
; Sequence 3, Application US/09053021  
; Patent No. 6270785  
; GENERAL INFORMATION:  
; APPLICANT: SELISKO, Barbara  
; APPLICANT: GARCIA-RODRIGUEZ, Consuelo  
; APPLICANT: ZAMUDIO-ZUNIGA, Fernando  
; APPLICANT: BECERRIL-LUJAN, Baltazar  
; APPLICANT: POSSANI-POSTAY, Lourival D.  
; TITLE OF INVENTION: Primary Sequence and cDNA of  
; Patent No. 6270785  
; TITLE OF INVENTION: Insecticidally Effective Toxins from Scorpions of the  
; TITLE OF INVENTION: Genus Centruroides  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSER: Dunner, L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/053,021  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/017,007  
; FILING DATE: 30-APR-1996  
; PRIOR APPLICATION NUMBER: US 08/848,261  
; APPLICATION NUMBER: US 08/848,261  
; FILING DATE: 29-APR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Garrett, Arthur S.  
; REGISTRATION NUMBER: 20,338  
; REFERENCE/DOCKET NUMBER: 06899.0001-01000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:



```
; ORGANISM: Hottentotta judaica
US-09-599-632-11

Query Match      18.2%; Score 35; DB 3; Length 243;
Best Local Similarity 56.5%; Pred. No. 0.051;
Matches 65; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 77 GCGATGACATGGAGACACGAGGAGTAATTCAGGCTATTGCTACGGAATTTGGTGCT 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 GTGATGGTATATGTAAGAGAAAGGAGGATATGCTATTGTTATTTCCCTGCTGTT 175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 137 ATTGCGAAGATGTTCCGATAGTACACCGACTTGGCCCCCTTTCTTAATAAAGATG 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 ATTGCGAAGATGAGAGATTAATCAAACTTTGGATAGAAAACCTAATAATG 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; PATENT NO. 6503729
; FILE REFERENCE: P8275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
```

```
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

Query Match 18.2%; Score 35; DB 3; Length 1664976;
Best Local Similarity 56.5%; Pred. No. 1.7;
Matches 65; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 3 AGACGGTTATCTGGTGGACAAGCGGCTGCAATACACTTGTGGATATTGGGAGAAA 62
DB 210409 ATACAGTTGTGGATTACAGCTTACCTGCAAGTTTATAGCTGGATATTGTGGAAGA 210468

QY 63 CAATATCTGCAATAGGGAATGCATGGAAGCCGAGGAGGTAAATTACGGCTAT 117
DB 210469 TAAGCCAGAAATGACATTTTATATGGAAGTGTCTTAGCTATAATTTCAGGTTT 210523

RESULT 5
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
```

```
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
```

LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (741684)..(741684)

OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g

Query Match 18.2%; Score 35; DB 3; Length 1664976;

Best Local Similarity 56.5%; Pred.No. 1.7;

Matches 65; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 3 AGACGGTTATCTGTGGCAAGACGGGCTGCAATACACTTGTGGATATTGGGAGAAA 62  
Db 210409 ATACAGTTGTGGATTAACAGCTTACCTGCAAGTTTATAGCTGGATATTGGAAGA 210468  
QY 63 CAAATACTGCAATAGGGAATGCACATGGAAGCAACCGAGGAGGTANTACGGCTAT 117  
Db 210469 TRAGCCCAAGATGACATTTTATATGGAAGTGTCTTAGCTATAATTTTCAGGTTT 210523

RESULT 6

US-09-830-230A-629  
; Sequence 629, Application US/09830230A  
; Patent No. 6902893  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Lyme Disease Vaccines  
; FILE REFERENCE: PB481US  
; CURRENT APPLICATION NUMBER: US/09/830,230A  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: PCT/US98/12718  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/057,483  
; PRIOR FILING DATE: 1997-09-03

```
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 629
; LENGTH: 3990
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1139)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1143)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1148)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1210)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1244)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1247)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1250)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1251)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-830-230A-629

Query Match
Best Local Similarity 16.9%; Score 32.4; DB 3; Length 3990;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 29 GCTGCAATACACTTGGTGGATATTGGGAGAAAACAATACTGCAATAGGGAATGCACAT 88
Db 361 GCTACAAATCCGATTGCTGCTATTGGGAATAAAGATGAGGATGCGGATTTTGGTGAT 420

Qy 89 GGAAGCACCGAGGAGTAATTACGGCTATTGCTACGGATTTGGTGCTATTGCCAAGGAT 148
Db 421 GGGATGAAGAAGGATGATCAGATTGCTGCTGCTATTGCTTTGAGGGGGATGGCTAAGGAT 480

Qy 149 TG 150
Db 481 GG 482

RESULT 7
US-09-125-619-3
; Sequence 3, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234USD3
; CURRENT APPLICATION NUMBER: US/09125619
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 629
; LENGTH: 3990
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1139)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1143)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1148)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1210)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1244)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1247)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1250)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1251)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-830-230A-629
```

```
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09125619
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 7766
; TYPE: DNA
; ORGANISM: Borrelia burgdorferi
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (127)
; OTHER INFORMATION: R = A OR G
; US-09-125-619-3

Query Match
Best Local Similarity 16.9%; Score 32.4; DB 3; Length 7766;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 29 GCTGCAATACACTTGGTGGATATTGGGAGAAAACAATACTGCAATAGGGAATGCACAT 88
Db 3979 GCTACAAATCCGATTGCTGCTATTGGGAATAAAGATGAGGATGCGGATTTTGGTGAT 4038

Qy 89 GGAAGCACCGAGGAGTAATTACGGCTATTGCTACGGATTTGGTGCTATTGCCAAGGAT 148
Db 4039 GGGATGAAGAAGGATGATCAGATTGCTGCTGCTATTGCTTTGAGGGGGATGGCTAAGGAT 4098

Qy 149 TG 150
Db 4099 GG 4100

RESULT 8
US-10-222-566-3
; Sequence 3, Application US/10222566
; Patent No. 6719983
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA
; FILE REFERENCE: UTSH:234USD3
; CURRENT APPLICATION NUMBER: US/10222566
; PRIOR FILING DATE: 2002-08-16
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 7766
; TYPE: DNA
; ORGANISM: Borrelia burgdorferi
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (127)
; OTHER INFORMATION: R = A OR G
; US-10-222-566-3

Query Match
Best Local Similarity 16.9%; Score 32.4; DB 3; Length 7766;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 29 GCTGCAATACACTTGGTGGATATTGGGAGAAAACAATACTGCAATAGGGAATGCACAT 88
Db 3979 GCTACAAATCCGATTGCTGCTATTGGGAATAAAGATGAGGATGCGGATTTTGGTGAT 4038

Qy 89 GGAAGCACCGAGGAGTAATTACGGCTATTGCTACGGATTTGGTGCTATTGCCAAGGAT 148
Db 4039 GGGATGAAGAAGGATGATCAGATTGCTGCTGCTATTGCTTTGAGGGGGATGGCTAAGGAT 4098

Qy 149 TG 150
Db 4099 GG 4100
```

Db 4039 GCGATGAGAGGATGATCAGATTGCTGCTGCTATTGCTTTGAGGGGATGCTAAGGAT 4098  
QY 149 TG 150  
Db 4099 GG 4100

RESULT 9

US-10-143-024A-3  
; Sequence 3, Application US/10143024A  
; Patent No. 6740744  
; GENERAL INFORMATION:  
; APPLICANT: NORRIS, STEVEN J.  
; APPLICANT: JING-REN, ZHANG  
; APPLICANT: HARDHAM, JOHN M.  
; APPLICANT: HOWELL, JERRILYN K.  
; APPLICANT: BARBOUR, ALAN G.  
; APPLICANT: WEINSTOCK, GEORGE M.  
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA  
; FILE REFERENCE: UTSH:234USD1  
; CURRENT APPLICATION NUMBER: US/10/143,024A  
; CURRENT FILING DATE: 2002-08-23  
; PRIOR APPLICATION NUMBER: 09/125,619  
; PRIOR FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: PCT/US97/02952  
; PRIOR FILING DATE: 1997-02-20  
; PRIOR APPLICATION NUMBER: 60/012,028  
; PRIOR FILING DATE: 1996-02-21  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 7766  
; TYPE: DNA  
; ORGANISM: Borrelia burgdorferi  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (127)  
; OTHER INFORMATION: R = A OR G  
US-10-143-024A-3

Query Match 16.9%; Score 32.4; DB 3; Length 7766;  
Best Local Similarity 54.1%; Pred. No. 1.5;  
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
QY 29 GCTGCAATACACTTCTGCTGATATTGGGAGAAACAAATCTGCAATAGGGAATGCACAT 88  
Db 3979 GCTACAAATCCGATTGCTGCTGCTATTGGGAATAAAGATGAGGATCGGATTTGGTGAT 4038  
QY 89 GGAAGCACCGAGGAGTAATACGGCTATTGCTACGGATTGGGTGCTATTGCGAAGGAT 148  
Db 4039 GGGATGAAGAAGGATGATCAGATTGCTGCTGCTATTGCTTTGAGGGGATGGCTAAGGAT 4098  
QY 149 TG 150  
Db 4099 GG 4100

RESULT 10

US-10-222-162-3  
; Sequence 3, Application US/10222162  
; Patent No. 6878616  
; GENERAL INFORMATION:  
; APPLICANT: NORRIS, STEVEN J.  
; APPLICANT: JING-REN, ZHANG  
; APPLICANT: HARDHAM, JOHN M.  
; APPLICANT: HOWELL, JERRILYN K.  
; APPLICANT: BARBOUR, ALAN G.  
; APPLICANT: WEINSTOCK, GEORGE M.  
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA  
; FILE REFERENCE: UTSH:234USD4  
; CURRENT APPLICATION NUMBER: US/10/222,162  
; CURRENT FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 09/125,619

; PRIOR FILING DATE: 1999-01-27  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 7766  
; TYPE: DNA  
; ORGANISM: Borrelia burgdorferi  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (127)  
; OTHER INFORMATION: R = A OR G  
US-10-222-162-3

Query Match 16.9%; Score 32.4; DB 3; Length 7766;  
Best Local Similarity 54.1%; Pred. No. 1.5;  
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
QY 29 GCTGCAATACACTTCTGCTGATATTGGGAGAAACAAATCTGCAATAGGGAATGCACAT 88  
Db 3979 GCTACAAATCCGATTGCTGCTGCTATTGGGAATAAAGATGAGGATCGGATTTGGTGAT 4038  
QY 89 GGAAGCACCGAGGAGTAATACGGCTATTGCTACGGATTGGGTGCTATTGCGAAGGAT 148  
Db 4039 GGGATGAAGAAGGATGATCAGATTGCTGCTGCTATTGCTTTGAGGGGATGGCTAAGGAT 4098  
QY 149 TG 150  
Db 4099 GG 4100

RESULT 11

US-09-949-016-17726/c  
; Sequence 17726, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17726  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17726

Query Match 16.7%; Score 32; DB 3; Length 601;  
Best Local Similarity 52.2%; Pred. No. 0.75;  
Matches 71; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 38 ACACTTGCTGGATATTGGGAGAAACAAATCTGCAATAGGGAATGCACATGGAGCACC 97  
Db 190 AGACAGACTGAATTGCGATGAAATATTTTGGAGGGAGGATGTAAATAGCCGCACA 131  
QY 98 GAGGAGTAATACGGCTATTGCTACGGATTGGGTGCTATTGCGAAGGATTTGCCGATA 157  
Db 130 AAGGGGTCCAACAGCTCTTTGAGTAGGCAATTTGGGGGTGTGTGTGGGG 71  
QY 158 GTACACCGACTTGGCC 173  
Db 70 GTGACCGAATTTGGC 55  
RESULT 12

US-09-949-016-17727/c  
; Sequence 17727, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

```

, , TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
, , FILE REFERENCE: CLO01307
, , CURRENT APPLICATION NUMBER: US/09/949,016
, , CURRENT FILING DATE: 2000-04-14
, , PRIOR APPLICATION NUMBER: 60/241,755
, , PRIOR FILING DATE: 2000-10-20
, , PRIOR APPLICATION NUMBER: 60/237,768
, , PRIOR FILING DATE: 2000-10-03
, , PRIOR APPLICATION NUMBER: 60/231,498
, , PRIOR FILING DATE: 2000-09-08
, , NUMBER OF SEQ ID NOS: 207012
, , SOFTWARE: FastSeq for Windows Version 4.0
, , SEQ ID NO 17727

```

	Query Match	16.7%	Score 32	DB 3	Length 601
	Best Local Similarity	52.2%	Pred. No. 0.75		
	Matches 71	Conservative 0	Mismatches 65	Indels 0	Gaps 0
Qy	38	ACACTTGTGGATATTGGGAAACAAATACTCAATAGGGAATGCACATAGGAAGCACC	97		
Db	266	AGACAGACTGNAATTTGGCATGAAATATATTTTGGGGAGGATGTAAATAGCCGCACA	207		
Qy	98	GAGGAGGTATATTACGGCTATTGCTACGGATTTGGGTGCTATTGGCAAGGATTTGCCGATA	157		
Db	206	AAGGGGTCACACAGCTCTTTGAGTAGCAATTTGTTAGAGCTTGGGGGTGTGTGTGTGGGG	147		
Qy	158	GTACACCGACTTGGCC	173		
Db	146	GTGACCGGAATTTGGC	131		

RESULT 13  
US-09-949-016-46726/c  
; Sequence 46726, Application US/09949016  
; Patent NO. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: Fast-SEO for Windows Version 4.0

	Query Match	16.7%	Score 32;	DB 3;	Length 601;
	Best Local Similarity	52.2%;	Pred. No. 0.75;		
	Matches 71;	Conservative 0;	Mismatches 65;	Indels 0;	Gaps 0;
QY	38	ACACTTCTCGATATTGGGAGAAACAAATAC	CTGCAATAGGGAATGCACATGGAAGCACC	97	
Db	190	AGACAGACTGAATTTGGCATGAAATATTTT	TTAGGAGGAGGATGTAAATAGCCGCACA	131	



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA: US/09/023,655  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1285:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2560 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: G312466  
US-09-023-655-1285

Query Match 16.7%; Score 32; DB 3; Length 2560;  
Best Local Similarity 52.2%; Pred. No. 1.3;  
Matches 71; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 38 ACACCTTGCTGGATATGGGAGAAACAATACTGCAATAGGGAATGCACATGGAAGCACC 97  
Db 2339 AGACAGACTGAATTTGCGATGAAATATTTTATAGAGGGAGGATGTAATAGCCGACA 2398  
QY 98 GAGGAGGTAATTACGGCTATTGCTACGGATTTGGGTGCTATTGCGAAGGATTGTCCGATA 157  
Db 2399 AAGGGGTCCACAGCTCTTTGATGAGCATTTGTAGACTTTGGGGGTGTGTGTGGGG 2458  
QY 158 GTACACCGACTTGGCC 173  
Db 2459 GTGGACCGAATTTGGC 2474

Search completed: January 22, 2006, 00:36:27  
Job time : 113 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:36:26 ; Search time 224 Seconds  
(without alignments)  
706.520 Million cell updates/sec

Title: US-10-721-793-115  
Perfect score: 192  
Sequence: 1 aaacaggttatctgggga.....cccttctaataaaagatgc 192

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6049916 seqs, 412136615 residues

Total number of hits satisfying chosen parameters: 12099832

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:

- 1: /cgn2\_6/ptodata/1/pubpna/US08 NEW PUB.seq:
- 2: /cgn2\_6/ptodata/1/pubpna/US06 NEW PUB.seq:
- 3: /cgn2\_6/ptodata/1/pubpna/US07 NEW PUB.seq:
- 4: /cgn2\_6/ptodata/1/pubpna/US09 NEW PUB.seq:
- 5: /cgn2\_6/ptodata/1/pubpna/US10 NEW PUB.seq:
- 6: /cgn2\_6/ptodata/1/pubpna/US11 NEW PUB.seq:
- 7: /cgn2\_6/ptodata/1/pubpna/US12 NEW PUB.seq:
- 8: /cgn2\_6/ptodata/1/pubpna/US13 NEW PUB.seq:
- 9: /cgn2\_6/ptodata/1/pubpna/US14 NEW PUB.seq:
- 10: /cgn2\_6/ptodata/1/pubpna/US15 NEW PUB.seq:
- 11: /cgn2\_6/ptodata/1/pubpna/US16 NEW PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30.8	16.0	3493	7	US-10-750-185-56579 Sequence 56579, A
2	30.8	16.0	3493	7	US-10-750-623-56579 Sequence 56579, A
3	30.2	15.7	1100	7	US-10-793-626-1201 Sequence 1201, Ap
4	30.2	15.7	3153	7	US-10-793-626-3441 Sequence 3441, Ap
5	29.8	15.5	2515	7	US-10-750-185-52777 Sequence 52777, A
6	29.8	15.5	2515	7	US-10-750-623-52777 Sequence 52777, A
7	29.6	15.4	4103	7	US-10-793-626-3532 Sequence 3532, Ap
8	29.2	15.2	3394	7	US-10-793-626-3629 Sequence 3629, Ap
9	29	15.1	1232	7	US-10-750-185-31353 Sequence 31353, A
10	29	15.1	1232	7	US-10-750-623-31353 Sequence 31353, A
11	28.8	15.0	1764	7	US-10-750-185-41036 Sequence 41036, A
12	28.8	15.0	1764	7	US-10-750-623-41036 Sequence 41036, A
13	28.2	14.7	1002	7	US-10-750-185-54668 Sequence 54668, A
14	28.2	14.7	1002	7	US-10-750-623-54668 Sequence 54668, A
15	28	14.6	1272	7	US-10-750-185-57766 Sequence 57766, A
16	28	14.6	1272	7	US-10-750-623-57766 Sequence 57766, A
17	28	14.6	175023	8	US-11-421-086-18 Sequence 18, Appl
18	27.8	14.5	2139	7	US-10-821-234-596 Sequence 596, App
19	27.8	14.5	2529	8	US-11-136-527-3182 Sequence 3182, Ap
20	27.8	14.5	24774	7	US-10-829-826B-53 Sequence 53, Appl
21	27.8	14.5	28920	7	US-10-829-826B-46 Sequence 46, Appl
22	27.8	14.5	29206	7	US-10-829-826B-56 Sequence 56, Appl

ALIGNMENTS

RESULT 1

US-10-750-185-56579  
; Sequence 56579, Application US/10750185  
; Publication No. US20050260603A1

; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen

; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 56579

; LENGTH: 3493

; TYPE: DNA

; ORGANISM: Bovine 19866880937875

US-10-750-185-56579

Query Match 16.0%; Score 30.8; DB 7; Length 3493;

Best Local Similarity 52.3%; Pred. No. 1;

Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 30 CTGCAATACATTGCTGGATATTGGGAGAAAACAATCTGCAATAGGGAATGCATG 89

Db 778 CTTAAAAACAATTGATCATCTTTAAAGAGGAAATATCTCAATTTTGCATG 837

QY 90 GAAGACCCGAGGAGTAAATTCAGCTATTGCTACGATTTGGTGTCTATTGCGAAGATT 149

Db 838 CTTAGAAAGAAACACATACATATATATACATGTTAGTTCAGTCGCTCACTG 897

QY 150 GTCCGATAGT 159

Db 898 GTCCGACTCT 907

RESULT 2

US-10-750-623-56579



Query Match 15.5%; Score 29.8; DB 7; Length 2515;  
Best Local Similarity 60.5%; Pred. No. 2;  
Matches 49; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
QY 3 AGACGGTTATCTGGTGGACAGACGGCTGCAATACACTTCTCGATATTTGGGAGAAA 62  
DB 121 AGATATTTCTCTGGTAGGAACTGTGAAGGAGTTCAAGGTCAGATCTTGAGAGAGAA 180  
QY 63 CAATATCTGCAATAGGAATG 83  
DB 181 ATCAATCTGCTATAAGCAATG 201

RESULT 6  
US-10-750-623-52777  
; Sequence 52777, Application US/10750623  
; Publication No. US20050287531A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-1  
; CURRENT APPLICATION NUMBER: US/10/750,623  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 52777  
; LENGTH: 2515  
; TYPE: DNA  
; ORGANISM: Bovine 19866880359186  
US-10-750-623-52777

Query Match 15.5%; Score 29.8; DB 7; Length 2515;  
Best Local Similarity 60.5%; Pred. No. 2;  
Matches 49; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
QY 3 AGACGGTTATCTGGTGGACAGACGGCTGCAATACACTTCTCGATATTTGGGAGAAA 62  
DB 121 AGATATTTCTCTGGTAGGAACTGTGAAGGAGTTCAAGGTCAGATCTTGAGAGAGAA 180  
QY 63 CAATATCTGCAATAGGAATG 83  
DB 181 ATCAATCTGCTATAAGCAATG 201

RESULT 7  
US-10-793-626-3532  
; Sequence 3532, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3532  
; LENGTH: 4103  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic

OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3532  
Query Match 15.4%; Score 29.6; DB 7; Length 4103;  
Best Local Similarity 50.7%; Pred. No. 2.9;  
Matches 71; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 52 TTGGAGAGAAAACAATACTGCAATAGGGAATGCAATGCAAGACACCGAGAGGTAAATTAC 111  
DB 3963 TTATGATAAATATAAATAATTAATGCCGATATTCATTGGAGGCCCGGAATTTGTAAGAAA 4022  
QY 112 GCGTATCTGCGAATTTGGTGGCTATTGCGAAGGATTTGTCGATAGTACACCGACTTGG 171  
DB 4023 AACACTTAACAAAAATTTAGGTGTTAATCTCGAGTATTACGCTAGTAGATTTTACTCG 4082  
QY 172 CCCCTTTCTAATAAAGATG 191  
DB 4083 ATTTGAAAAAATGATAGATG 4102

RESULT 8  
US-10-793-626-3629/c  
; Sequence 3629, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3629  
; LENGTH: 3394  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3629

Query Match 15.2%; Score 29.2; DB 7; Length 3394;  
Best Local Similarity 53.5%; Pred. No. 3.7;  
Matches 61; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 29 GCTGCAATACACTTCTGCGATATTTGGGAGAAAACAATACTGCAATAGGAGATGCACAT 88  
DB 2110 GCAGCGAACTTAGAGCGGCATATATGCCGAATGCTGATAAGGTATTAGACATTCATCT 2051  
QY 89 GGAAGCACCGAGGAGTAAATGCGCTATTGCTACCGATTGGGTGCTATTGG 142  
DB 2050 GAATTAATGATGATTTAAAAAAGGCTAACGCATAGGAGGTGATCACATGAG 1997

RESULT 9  
US-10-750-185-31353  
; Sequence 31353, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482

```
; PRIORITY FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31353
; LENGTH: 1232
; TYPE: DNA
; ORGANISM: Bovine 19866880618861
US-10-750-185-31353

Query Match      15.1%; Score 29; DB 7; Length 1232;
Best Local Similarity 58.8%; Pred. No. 2.9;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 23 AGACGGGCTGCAATACACTTCTGATATTTGGGAGAAAACAATACATGCAATAGGGAAT 82
Db 255 AAAGGGGAGAAATAACAAATGTTGGAGAAAAGGAACCTTCATATCTGTTGGTGGGAAT 314
QY 83 GCACATGGAAGCAGCAGGAGGTAA 107
Db 315 GTAAATTGATGAGCCACTATGGAA 339

RESULT 10
US-10-750-623-31353
; Sequence 31353, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31353
; LENGTH: 1232
; TYPE: DNA
; ORGANISM: Bovine 19866880618861
US-10-750-623-31353

Query Match      15.1%; Score 29; DB 7; Length 1232;
Best Local Similarity 58.8%; Pred. No. 2.9;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 23 AGACGGGCTGCAATACACTTCTGATATTTGGGAGAAAACAATACATGCAATAGGGAAT 82
Db 255 AAAGGGGAGAAATAACAAATGTTGGAGAAAAGGAACCTTCATATCTGTTGGTGGGAAT 314
QY 83 GCACATGGAAGCAGCAGGAGGTAA 107
Db 315 GTAAATTGATGAGCCACTATGGAA 339

RESULT 11
US-10-750-185-41036
; Sequence 41036, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 41036
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Bovine 19866881141918
US-10-750-185-41036

Query Match      15.0%; Score 28.8; DB 7; Length 1764;
Best Local Similarity 54.8%; Pred. No. 3.9;
Matches 57; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 9 TTATCTGCTGGACAAAGACGGGCTGCAATACACTTGTGGATATTTGGGAGAAAACAATA 68
Db 1242 TTGCTTGGGAATCAGTGGGGTTGCAAGAGTTCAGACGTGACTTGGCAACTAAACAACA 1301
QY 69 CTGCAATAGGAATGCATGGAAGCACCGAGGAGGTAAATTACG 112
Db 1302 CAACAAAAATGATACACGTGAATGGCCCGGCGAGGAATTCAG 1345

RESULT 12
US-10-750-623-41036
; Sequence 41036, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 41036
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Bovine 19866881141918
US-10-750-623-41036

Query Match      15.0%; Score 28.8; DB 7; Length 1764;
Best Local Similarity 54.8%; Pred. No. 3.9;
Matches 57; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 9 TTATCTGCTGGACAAAGACGGGCTGCAATACACTTGTGGATATTTGGGAGAAAACAATA 68
Db 1242 TTGCTTGGGAATCAGTGGGGTTGCAAGAGTTCAGACGTGACTTGGCAACTAAACAACA 1301
QY 69 CTGCAATAGGAATGCATGGAAGCACCGAGGAGGTAAATTACG 112
Db 1302 CAACAAAAATGATACACGTGAATGGCCCGGCGAGGAATTCAG 1345

RESULT 13
US-10-750-185-54668
; Sequence 54668, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 54668
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Bovine 19866881141918
US-10-750-185-54668

Query Match      15.0%; Score 28.8; DB 7; Length 1764;
Best Local Similarity 54.8%; Pred. No. 3.9;
Matches 57; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 9 TTATCTGCTGGACAAAGACGGGCTGCAATACACTTGTGGATATTTGGGAGAAAACAATA 68
Db 1242 TTGCTTGGGAATCAGTGGGGTTGCAAGAGTTCAGACGTGACTTGGCAACTAAACAACA 1301
QY 69 CTGCAATAGGAATGCATGGAAGCACCGAGGAGGTAAATTACG 112
Db 1302 CAACAAAAATGATACACGTGAATGGCCCGGCGAGGAATTCAG 1345
```

; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 54668  
; LENGTH: 1002  
; TYPE: DNA  
; ORGANISM: Bovine 19866880898655  
US-10-750-185-54668

Query Match 14.7%; Score 28.2; DB 7; Length 1002;  
Best Local Similarity 61.6%; Pred. No. 5;  
Matches 45; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
  
QY 75 TAGGGAATGCACATGGAAGCAGCGAGGTAATTACGGCTATTGCTACGGATTTGGGTG 134  
Db 575 TGGAGATGCTCATGATACAGAAAGCAGGATAGTGCTATGACACAGTATCTGGCTA 634  
  
QY 135 CTATTGCGAAGGA 147  
Db 635 CTATGTGTTAGGA 647

RESULT 14  
US-10-750-623-54668  
; Sequence 54668, Application US/10750623  
; Publication No. US20050287531A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-1  
; CURRENT APPLICATION NUMBER: US/10/750,623  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 54668  
; LENGTH: 1002  
; TYPE: DNA  
; ORGANISM: Bovine 19866880898655  
US-10-750-623-54668

Query Match 14.7%; Score 28.2; DB 7; Length 1002;  
Best Local Similarity 61.6%; Pred. No. 5;  
Matches 45; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
  
QY 75 TAGGGAATGCACATGGAAGCAGCGAGGTAATTACGGCTATTGCTACGGATTTGGGTG 134  
Db 575 TGGAGATGCTCATGATACAGAAAGCAGGATAGTGCTATGACACAGTATCTGGCTA 634  
  
QY 135 CTATTGCGAAGGA 147  
Db 635 CTATGTGTTAGGA 647

RESULT 15  
US-10-750-185-57766

; Sequence 57766, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 57766  
; LENGTH: 1272  
; TYPE: DNA  
; ORGANISM: Bovine 19866881046656  
US-10-750-185-57766

Query Match 14.6%; Score 28; DB 7; Length 1272;  
Best Local Similarity 66.7%; Pred. No. 6.5;  
Matches 40; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
  
QY 77 GGGAAATGCACATGGAAGCAGCGAGGTAATTACGGCTATTGCTACGGATTTGGGTGCT 136  
Db 1119 GAGAAAGCGAGATGGAAGGACTGTGAAGGTGACTATAGGACTTGGTGGGTGTT 1178

Search completed: January 22, 2006, 01:16:42  
Job time : 225 secs

*This Page Blank (usps)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:31:18 ; Search time 2178 Seconds  
(without alignments)  
728.981 Million cell updates/sec

Title: US-10-721-793-115  
Perfect score: 192  
Sequence: 1 aaagacggttctcgttga.....cccttctaataaagatgc 192

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	192	9	US-10-721-793-115
2	192	100.0	323	9	US-10-721-793-113
3	190.4	99.2	192	9	US-10-721-793-119
4	190.4	99.2	323	9	US-10-721-793-117
5	156.8	81.7	192	9	US-10-721-793-151
6	156.8	81.7	320	9	US-10-721-793-149
7	155.2	80.8	192	9	US-10-721-793-195
8	155.2	80.8	320	9	US-10-721-793-193
9	152	79.2	192	9	US-10-721-793-47
10	152	79.2	258	9	US-10-721-793-45
11	150.4	78.3	192	9	US-10-721-793-43
12	150.4	78.3	254	9	US-10-721-793-41
13	145.4	75.7	198	9	US-10-721-793-91
14	145.4	75.7	323	9	US-10-721-793-89
15	126.8	66.0	189	9	US-10-721-793-111
16	126.8	66.0	311	9	US-10-721-793-109
17	124.6	64.9	323	9	US-10-721-793-101
18	124.6	64.9	323	9	US-10-721-793-105
19	123.2	64.2	192	9	US-10-721-793-103
20	123.2	64.2	192	9	US-10-721-793-107
21	123	64.1	195	9	US-10-721-793-147
22	123	64.1	198	9	US-10-721-793-175
23	123	64.1	323	9	US-10-721-793-145

24	123	64.1	323	9	US-10-721-793-173	Sequence 173, App
25	119.8	62.4	198	9	US-10-721-793-67	Sequence 67, App
26	119.8	62.4	198	9	US-10-721-793-71	Sequence 71, Appl
27	119.8	62.4	198	9	US-10-721-793-159	Sequence 159, App
28	119.8	62.4	198	9	US-10-721-793-163	Sequence 163, App
29	119.8	62.4	322	9	US-10-721-793-65	Sequence 65, Appl
30	119.8	62.4	322	9	US-10-721-793-69	Sequence 69, Appl
31	119.8	62.4	323	9	US-10-721-793-157	Sequence 157, App
32	119.8	62.4	323	9	US-10-721-793-139	Sequence 139, App
33	118.2	61.6	195	9	US-10-721-793-139	Sequence 139, App
34	118.2	61.6	198	9	US-10-721-793-155	Sequence 155, App
35	118.2	61.6	319	9	US-10-721-793-153	Sequence 153, App
36	118.2	61.6	323	9	US-10-721-793-137	Sequence 137, App
37	116.6	60.7	195	9	US-10-721-793-83	Sequence 83, Appl
38	116.6	60.7	274	9	US-10-721-793-81	Sequence 81, Appl
39	116.6	60.7	323	9	US-10-721-793-97	Sequence 97, Appl
40	115	59.9	195	9	US-10-721-793-167	Sequence 167, App
41	115	59.9	195	9	US-10-721-793-171	Sequence 171, App
42	115	59.9	323	9	US-10-721-793-165	Sequence 165, App
43	115	59.9	323	9	US-10-721-793-169	Sequence 169, App
44	113.4	59.1	195	9	US-10-721-793-23	Sequence 23, Appl
45	113.4	59.1	195	9	US-10-721-793-143	Sequence 143, App

ALIGNMENTS

RESULT 1  
US-10-721-793-115  
; Sequence 115, Application US/10721793  
; Publication No. US20050065331A1  
; GENERAL INFORMATION:  
; APPLICANT: Corona Villegas, Miguel  
; APPLICANT: Garcia Rodriguez, Ma Consuelo  
; APPLICANT: Valdez Cruz, Norma Adriana  
; APPLICANT: Gurrula Briones, Georgina  
; APPLICANT: Berceril Lujan, Baltazar  
; APPLICANT: Possani Postay, Lourival Domingos  
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the  
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides  
; FILE REFERENCE: 2099.0070001  
; CURRENT APPLICATION NUMBER: US/10721,793  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: US 60/430,067  
; PRIOR FILING DATE: 2002-12-02  
; NUMBER OF SEQ ID NOS: 294  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 115  
; LENGTH: 192  
; TYPE: DNA  
; ORGANISM: Centruroides elegans  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(192)  
; OTHER INFORMATION: Product= Sodium-channel modifier toxin  
US-10-721-793-115

Query Match	100.0%;	Score 192;	DB 9;	Length 192;
Best Local Similarity	100.0%;	Pred. No. 1.1e-53;		
Matches 192;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AAAGACGGTTATCTGGTGACAAAGACGGGCTGCAAAATACACTTGTGGATTTGGGAGAA	60	
Db	1	AAAGACGGTTATCTGGTGACAAAGACGGGCTGCAAAATACACTTGTGGATTTGGGAGAA	60	
QY	61	AACAAATCTGCAATAGGGAATGCATGGAAGCAGCAGGAGGTATTAACGGCTATTGC	120	
Db	61	AACAAATCTGCAATAGGGAATGCATGGAAGCAGCAGGAGGTATTAACGGCTATTGC	120	
QY	121	TACGGATTTGGTGCTATTGGGAAGGATTTGCGATAGTACACCGACTTGGCCCTTTCT	180	
Db	121	TACGGATTTGGTGCTATTGGGAAGGATTTGCGATAGTACACCGACTTGGCCCTTTCT	180	

Qy 181 AATAAAGATGC 192  
|||||  
Db 181 AATAAAGATGC 192

## RESULT 2

US-10-721-793-113  
; Sequence 113, Application US/10721793  
; Publication No. US20050065331A1  
; GENERAL INFORMATION:  
; APPLICANT: Corona Villegas, Miguel  
; APPLICANT: Garcia Rodriguez, Ma Consuelo  
; APPLICANT: Valdez Cruz, Norma Adriana  
; APPLICANT: Gurrola Briones, Georgina  
; APPLICANT: Becerril Lujan, Baltazar  
; APPLICANT: Possani Postay, Lourival Domingos  
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the  
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides  
; FILE REFERENCE: 2099.0070001  
; CURRENT APPLICATION NUMBER: US/10/721,793  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: US 60/430,067  
; PRIOR FILING DATE: 2002-12-02  
; NUMBER OF SEQ ID NOS: 294  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 113  
; LENGTH: 323  
; TYPE: DNA  
; ORGANISM: Centruroides elegans  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (5)..(265)  
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor  
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly  
; OTHER INFORMATION: and the last 2 basic aminoacids are cut  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (269)..(323)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(4)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (65)..( )  
; OTHER INFORMATION: Product= Sodium-channel modifier toxin  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (5)..(64)  
; OTHER INFORMATION:  
US-10-721-793-113

Query Match 100.0%; Score 192; DB 9; Length 323;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAAGACGGTTATCTGGTGACAAAGACGGGCTGCAAATACACTTGTCTGGATATTGGAGAA 60  
Db 65 AAAGACGGTTATCTGGTGACAAAGACGGGCTGCAAATACACTTGTCTGGATATTGGAGAA 124  
Qy 61 AACAAATCTGCAATAGGGAATGCATGGAAGCACCGAGAGGTAATTACGGCTATTGC 120  
Db 125 AACAAATCTGCAATAGGGAATGCATGGAAGCACCGAGAGGTAATTACGGCTATTGC 184  
Qy 121 TACGGATTGGTGCTATTGGCAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180  
Db 185 TACGGATTGGTGCTATTGGCAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 244  
Qy 181 AATAAAGATGC 192  
|||||  
Db 245 AATAAAGATGC 256

## RESULT 3

US-10-721-793-119  
; Sequence 119, Application US/10721793  
; Publication No. US20050065331A1  
; GENERAL INFORMATION:  
; APPLICANT: Corona Villegas, Miguel  
; APPLICANT: Garcia Rodriguez, Ma Consuelo  
; APPLICANT: Valdez Cruz, Norma Adriana  
; APPLICANT: Gurrola Briones, Georgina  
; APPLICANT: Becerril Lujan, Baltazar  
; APPLICANT: Possani Postay, Lourival Domingos  
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the  
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides  
; FILE REFERENCE: 2099.0070001  
; CURRENT APPLICATION NUMBER: US/10/721,793  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: US 60/430,067  
; PRIOR FILING DATE: 2002-12-02  
; NUMBER OF SEQ ID NOS: 294  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 119  
; LENGTH: 192  
; TYPE: DNA  
; ORGANISM: Centruroides elegans  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(192)  
; OTHER INFORMATION: Product= Sodium-channel modifier toxin  
US-10-721-793-119

Query Match 99.2%; Score 190.4; DB 9; Length 192;  
Best Local Similarity 99.5%; Pred. No. 3.6e-53;  
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 AAAGACGGTTATCTGGTGACAAAGACGGGCTGCAAATACACTTGTCTGGATATTGGAGAA 60  
Db 1 AAAGACGGTTATCTGGTGACAAAGACGGGCTGCAAATACACTTGTCTGGATATTGGAGAA 60  
Qy 61 AACAAATCTGCAATAGGGAATGCATGGAAGCACCGAGAGGTAATTACGGCTATTGC 120  
Db 61 AACAAATCTGCAATAGGGAATGCATGGAAGCACCGAGAGGTAATTACGGCTATTGC 120  
Qy 121 TACGGATTGGTGCTATTGGCAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180  
Db 121 TACGGATTGGTGCTATTGGCAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180  
Qy 181 AATAAAGATGC 192  
|||||  
Db 181 AATAAAGATGC 192

## RESULT 4

US-10-721-793-117  
; Sequence 117, Application US/10721793  
; Publication No. US20050065331A1  
; GENERAL INFORMATION:  
; APPLICANT: Corona Villegas, Miguel  
; APPLICANT: Garcia Rodriguez, Ma Consuelo  
; APPLICANT: Valdez Cruz, Norma Adriana  
; APPLICANT: Gurrola Briones, Georgina  
; APPLICANT: Becerril Lujan, Baltazar  
; APPLICANT: Possani Postay, Lourival Domingos  
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the  
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides  
; FILE REFERENCE: 2099.0070001  
; CURRENT APPLICATION NUMBER: US/10/721,793  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: US 60/430,067  
; PRIOR FILING DATE: 2002-12-02  
; NUMBER OF SEQ ID NOS: 294  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 117

```
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Centruroides elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(265)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (269)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (65)..(1)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(64)
; OTHER INFORMATION:
US-10-721-793-117

Query Match          99.2%; Score 190.4; DB 9; Length 323;
Best Local Similarity 99.5%; Pred. No. 4.5e-53;
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAGCGGTTATCTGGTGACAAAGCGGGCTGCAATACACTTCTGGATATTGGAGAA 60
DB 65 AAAGCGGTTATCTGGTGACAAAGCGGGCTGCAATACACTTCTGGATATTGGAGAA 124
QY 61 AACAAATCTGCAATAGGGAATGCATCGAAGACCCGAGGAGGTAATTAACGGCTATTGC 120
DB 125 AACAAATCTGCAATAGGGAATGCATCGAAGACCCGAGGAGGTAATTAACGGCTATTGC 184
QY 121 TACGGATTTGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 180
DB 185 TACGGATTTGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 244
QY 181 AATAAAGATGC 192
DB 245 AATAAAGATGC 256

RESULT 5
US-10-721-793-151
; Sequence 151, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 151
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(262)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (62)..(1)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
```

```
; NAME/KEY: CDS
; LOCATION: (1)..(192)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(192)
US-10-721-793-151

Query Match          81.7%; Score 156.8; DB 9; Length 192;
Best Local Similarity 88.5%; Pred. No. 6.1e-42;
Matches 170; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 AAAGCGGTTATCTGGTGACAAAGCGGGCTGCAATACACTTCTGGATATTGGAGAA 60
DB 1 AAGGACGGTTATCTAGTGGAAAGACGGGCTGCAAAAGACTTGTCTACAAATTGGAGAA 60
QY 61 AACAAATCTGCAATAGGGAATGCATCGAAGACCCGAGGAGGTAATTAACGGCTATTGC 120
DB 61 AACGATTTTTCATATAGGGAATGCAATGGAAGCACATAGGAGGTAGTTATGGCTATTTC 120
QY 121 TACGGATTTGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 180
DB 121 TACGGATTTGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 180
QY 181 AATAAAGATGC 192
DB 181 AATAAAGATGC 192

RESULT 6
US-10-721-793-149
; Sequence 149, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 149
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(262)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (62)..(1)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
```

; LOCATION: (5)..(61)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (266)..(320)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(4)  
; OTHER INFORMATION:  
; PUBLICATION INFORMATION:  
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.  
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,  
; TITLE: that recognize Na<sup>+</sup>-channels  
; JOURNAL: Toxicon  
; VOLUME: 39  
; ISSUE: 12  
; PAGES: 1893-1898  
; DATE: 2001-12-01  
; DATABASE ENTRY DATE:  
; RELEVANT RESIDUES: (5)..(262)  
US-10-721-793-149

Query Match 81.7%; Score 156.8; DB 9; Length 320;  
Best Local Similarity 88.5%; Pred. No. 7,6e-42;  
Matches 170; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
  
QY 1 AAAGACGGTTATCTGTCGACAAAGACGGCTGCAAAATACACTTCTGGATATTGGAGAA 60  
DB 62 AAGGACGGTTATCTAGTGGAAAGACGGCTGCAAAAGACTTGTACAAATTGGAGAA 121  
  
QY 61 AACAAATCTGCAATAGGGAATGCACATGGAGCCAGGAGGTAATTACGGCTATTGC 120  
DB 122 AACGATTTTGCATAGGGAATGCAATGGAAAGCACATAGGAGGTAGTTATGGCTATTTC 181  
  
QY 121 TACGATTTGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180  
DB 182 TACGATTTGGTGCTATTGCGAAGGATTGCGGATAGTCCCGATAGTACAGACTTGGCCCTTCCT 241  
  
QY 181 AATAAAGATGC 192  
DB 242 AATAAAGATGC 253

RESULT 7  
US-10-721-793-195  
; Sequence 195, Application US/10721793  
; Publication No. US20050065331A1  
; GENERAL INFORMATION:  
; APPLICANT: Corona Villegas, Miguel  
; APPLICANT: Garcia Rodriguez, Ma Consuelo  
; APPLICANT: Valdez Cruz, Norma Adriana  
; APPLICANT: Gurrola Briones, Georgina  
; APPLICANT: Becerril Lujan, Baltazar  
; APPLICANT: Possani Postay, Lourival Domingos  
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the  
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides  
; FILE REFERENCE: 2099.0070001  
; CURRENT APPLICATION NUMBER: US/10/721.793  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: US 60/430,067  
; PRIOR FILING DATE: 2002-12-02  
; NUMBER OF SEQ ID NOS: 294  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 195  
; LENGTH: 192  
; TYPE: DNA  
; ORGANISM: Centruroides sculpturatus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(192)  
; OTHER INFORMATION: Product= Sodium-channel modifier toxin  
; OTHER INFORMATION: Information:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(4)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mat\_peptide

; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.  
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,  
; TITLE: that recognize Na<sup>+</sup>-channels  
; JOURNAL: Toxicon  
; VOLUME: 39  
; ISSUE: 12  
; PAGES: 1893-1898  
; DATE: 2001-12-01  
; DATABASE ENTRY DATE:  
; RELEVANT RESIDUES: (1)..(192)  
US-10-721-793-195

Query Match 80.8%; Score 155.2; DB 9; Length 192;  
Best Local Similarity 88.0%; Pred. No. 2.1e-41;  
Matches 169; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
  
QY 1 AAAGACGGTTATCTGTCGACAAAGACGGCTGCAAAATACACTTCTGGATATTGGAGAA 60  
DB 1 AAGGAAGTTTATCTGTCGACGTAAGGCTGCAAAAAAATTCCTGGAAATTGGAGAT 60  
  
QY 61 AACAAATCTGCAATAGGGAATGCACATGGAGCCAGGAGGTAATTACGGCTATTGC 120  
DB 61 AACGATTTTGCATAGGGAATGTAATGGAAAGCACATAGGAGGTAGTTACGGCTATTGC 120  
  
QY 121 TACGATTTGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180  
DB 121 TACGATTTGGTGCTATTGCGAAGGATTGCGGATAGTCCCGATAGTACAGACTTGGCCCTTCCT 180  
  
QY 181 AATAAAGATGC 192  
DB 181 AATAAAGATGC 192

RESULT 8  
US-10-721-793-193  
; Sequence 193, Application US/10721793  
; Publication No. US20050065331A1  
; GENERAL INFORMATION:  
; APPLICANT: Corona Villegas, Miguel  
; APPLICANT: Garcia Rodriguez, Ma Consuelo  
; APPLICANT: Valdez Cruz, Norma Adriana  
; APPLICANT: Gurrola Briones, Georgina  
; APPLICANT: Becerril Lujan, Baltazar  
; APPLICANT: Possani Postay, Lourival Domingos  
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the  
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides  
; FILE REFERENCE: 2099.0070001  
; CURRENT APPLICATION NUMBER: US/10/721.793  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: US 60/430,067  
; PRIOR FILING DATE: 2002-12-02  
; NUMBER OF SEQ ID NOS: 294  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 193  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Centruroides sculpturatus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (5)..(262)  
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor  
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly  
; OTHER INFORMATION: and the last 2 basic aminoacids are cut  
; FEATURE:  
; NAME/KEY: 5'clip  
; LOCATION: (266)..(320)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(4)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mat\_peptide

```
; LOCATION: (62)...()
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (5)...(61)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUES: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (5)...(262)
US-10-721-793-193

Query Match      80.8%; Score 155.2; DB 9; Length 320;
Best Local Similarity 88.0%; Pred. No. 2.6e-41;
Matches 169; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGGTGACACAGACGGGCTGCAAAATACACTTGTGGATATTGGAGAA 60
DB 62 AAGGAAGGTTATCTGGTGACACGTAAGGGCTGCAAAAAAATGCTGGAAATGGAGAT 121
QY 61 AACAAATACTGCAATAGGAATGCACATGGAAGACCGGAGGAGTAATTACGGCTATTGC 120
DB 122 AAGGATTATGCAATAGGAGATGTAATGGAAGCACATAGGAGGTAGTTACGGCTATTGC 181
QY 121 TACGGATTGGTGCTATTGCGAAGGATGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 182 TAGCGATTGGTGCTATTGCGAAGGATGTCGATAGTACACAGACTTGGCCCTTTCT 241
QY 181 AATAAAGATGC 192
DB 242 AATAAAGATGC 253

RESULT 9
US-10-721-793-47
; Sequence 47, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 47
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Centruroides exilicauda
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(192)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gl
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (205)...(258)
; OTHER INFORMATION:
US-10-721-793-47

Query Match      79.2%; Score 152; DB 9; Length 258;
Best Local Similarity 87.0%; Pred. No. 2.8e-40;
Matches 167; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGGTGACACAGACGGGCTGCAAAATACACTTGTGGATATTGGAGAA 60
DB 1 AAGGACGGTTATCTGGTGAGGTCACGGGCTGCAAAAAGTCTTGCTATAAATTTGGAGAA 60
QY 61 AACAAATACTGCAATAGGAGATGCACATGGAAGACCGGAGGAGTAATTACGGCTATTGC 120
DB 61 AACAAATCTGCAATAGGAGATGCAAAATGAAGCACCGGAGGAGTAGTTACGGCTATTGC 120
QY 121 TACGGATTGGTGCTATTGCGAAGGATGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 121 TATTTTTTTGGTGCTATTGCGAAGGATGTCGCGAAGGATGTCGCGAAGTACACCGACTTGGCCCTTTCT 180
QY 181 AATAAAGATGC 192
DB 181 AATAAATCATGC 192
```

```
QY 1 AAAGACGGTTATCTGGTGACACAGACGGGCTGCAAAATACACTTGTGGATATTGGAGAA 60
DB 1 AAGGACGGTTATCTGGTGAGGTCACGGGCTGCAAAAAGTCTTGCTATAAATTTGGAGAA 60
QY 61 AACAAATACTGCAATAGGAATGCACATGGAAGACCGGAGGAGTAATTACGGCTATTGC 120
DB 61 AACAAATCTGCAATAGGAGATGCAAAATGAAGCACCGGAGGAGTAGTTACGGCTATTGC 120
QY 121 TACGGATTGGTGCTATTGCGAAGGATGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 121 TATTTTTTTGGTGCTATTGCGAAGGATGTCGCGAAGGATGTCACCGACTTGGCCCTTTCT 180
QY 181 AATAAAGATGC 192
DB 181 AATAAATCATGC 192

RESULT 10
US-10-721-793-45
; Sequence 45, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 45
; LENGTH: 258
; TYPE: DNA
; ORGANISM: Centruroides exilicauda
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(204)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gl
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (205)...(258)
; OTHER INFORMATION:
US-10-721-793-45

Query Match      79.2%; Score 152; DB 9; Length 258;
Best Local Similarity 87.0%; Pred. No. 2.8e-40;
Matches 167; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGGTGACACAGACGGGCTGCAAAATACACTTGTGGATATTGGAGAA 60
DB 1 AAGGACGGTTATCTGGTGAGGTCACGGGCTGCAAAAAGTCTTGCTATAAATTTGGAGAA 60
QY 61 AACAAATACTGCAATAGGAGATGCACATGGAAGACCGGAGGAGTAATTACGGCTATTGC 120
DB 61 AACAAATCTGCAATAGGAGATGCAAAATGAAGCACCGGAGGAGTAGTTACGGCTATTGC 120
QY 121 TACGGATTGGTGCTATTGCGAAGGATGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 121 TATTTTTTTGGTGCTATTGCGAAGGATGTCGCGAAGGATGTCACCGACTTGGCCCTTTCT 180
QY 181 AATAAAGATGC 192
DB 181 AATAAATCATGC 192
```



```
Db      1  AAGGAGGTTATCTGCTGAACAAAGACACAGGCTGTAAATACAACTGCTTGATATTGGGA 60
Qy      58  GAAACAAATATCTGCAATAGGAATGCATGGAAGCACCAGGAGGTAAATTACGGCTAT 117
Db      61  GAAACAAACAACTGCGATATGGAATGCAAGCGAAGAACCAAGGAGGTAGTTACGGCTAT 120
Qy      118  TGCTACGGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTT 177
Db      121  TGCTACGGATTGGGTGCTATTGTAAGGATTGTCGATAGTACACCGACTTGGCCCTT 180
Qy      178  TCTAATAAAAGATGC 192
Db      181  CCTAATAAACATGC 195

RESULT 14
US-10-721-793-89
; Sequence 89, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Centruroides noxius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(265)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Ser is amidated, and the last Gly
; OTHER INFORMATION: and the last basic aminoacid are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (269)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; NAME/KEY: mat_peptide
; LOCATION: (62)..()
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
US-10-721-793-89

Query Match      75.7%; Score 145.4; DB 9; Length 323;
Best Local Similarity 87.7%; Pred. No. 4.9e-38;
Matches 171; Conservative 0; Mismatches 21; Indels 3; Gaps 1;

Qy      1  AAGAGCGTTATCTGGTGGACAAAGCACAGGCTGTAATACAACTGCTTGATATTGGGA 57
Db      62  AAGGAGGTTATCTGGTGAACAAAGCACAGGCTGTAATACAACTGCTTGATATTGGGA 121
Qy      58  GAAACAAATATCTGCAATAGGAATGCATGGAAGCACCAGGAGGTAAATTACGGCTAT 117
```

```
Db      122  GAAACAAACAACTGCGATATGGAATGCAAGCGAAGAACCAAGGAGGTAGTTACGGCTAT 181
Qy      118  TGCTACGGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTT 177
Db      182  TGCTACGGATTGGGTGCTATTGTAAGGATTGTCGATAGTACACCGACTTGGCCCTT 241
Qy      178  TCTAATAAAAGATGC 192
Db      242  CCTAATAAACATGC 256

RESULT 15
US-10-721-793-111
; Sequence 111, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 111
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Centruroides elegans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(189)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
US-10-721-793-111

Query Match      66.0%; Score 126.8; DB 9; Length 189;
Best Local Similarity 82.8%; Pred. No. 6.4e-32;
Matches 159; Conservative 0; Mismatches 27; Indels 6; Gaps 1;

Qy      1  AAGAGCGTTATCTGGTGGACAAAGCAGCGCTGCAATACACTGCTGGATATTGGGAGAA 60
Db      4  AAGGAGCGTTATCTGGTGAACAAAGCAGCGCTGCAATACAAATTGCTGGATATTGGGAGAA 63
Qy      61  AACAAATACTGCAATAGGGAATGCATATGGAAGCACCGGAGGTAAATTACGGCTATTGC 120
Db      64  AACAAATACTGCAATTCGGGAATGC-----AAAGAGGTAGGTGCTGTTACGGCTATTGC 117
Qy      121  TAGGGATTTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 180
Db      118  TAGCTCTTTGGGTGCTATTGCGAAGGATTACCGAAGCGTACTGACCTGGCCCTTCT 177
Qy      181  AATAAAAGATGC 192
Db      178  GATAAAACATGC 189

Search completed: January 22, 2006, 01:12:50
Job time : 2178 secs
```

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:29:01 ; Search time 8943 Seconds  
(without alignments)  
1004.486 Million cell updates/sec

Title: US-10-721-793-115  
Perfect score: 192  
Sequence: 1 aaagcgggtatctgtgga.....ccctttctaataaagatgc 192

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36.8	19.2	582	3	BP248044
2	35.2	18.3	656	2	BF640569
3	35	18.2	982	10	CNS000JOE
4	34.8	18.1	820	1	AU119061
5	34.8	18.1	849	8	DN839051
6	34.6	18.0	340	6	CP503342
7	34.6	18.0	918	5	BUS52525
8	34.4	17.9	1080	8	DR738811
9	34.2	17.8	459	11	CR899083
10	34.2	17.8	450	5	BQ864839
11	34.2	17.8	893	5	BUS43170
12	34	17.7	748	4	AY440472
13	33.8	17.6	704	8	DR423779
14	33.6	17.5	377	8	N75948
15	33.6	17.5	474	8	N57766
16	33.6	17.5	581	9	AQ725129
17	33.4	17.4	978	10	CL041476
18	33.4	17.4	1213	10	CL644174
19	33.4	17.4	1226	1	AG181747
20	33.2	17.3	352	1	AI468868
21	33.2	17.3	619	1	AA293053
22	33	17.2	336	2	BF767499

c	23	33	17.2	659	9	CC453499	CC453499
	24	33	17.2	761	9	CC847597	CC847597
	25	33	17.2	838	10	BX970645	BX970645
c	26	33	17.2	903	5	BX436789	BX436789
	27	33	17.2	937	2	BG742646	BG742646
c	28	33	17.2	1013	5	BX332631	BX332631
	29	32.8	17.1	246	5	BZ38563	BZ38563
c	30	32.8	17.1	631	2	BE901242	BE901242
	31	32.8	17.1	777	2	BE902136	BE902136
	32	32.6	17.1	951	2	BE742069	BE742069
	33	32.6	17.0	559	6	CP722903	CP722903
	34	32.6	17.0	564	6	CB285819	CB285819
	35	32.6	17.0	596	1	AJ657416	AJ657416
c	36	32.6	17.0	619	8	DN106755	DN106755
	37	32.6	17.0	620	8	DN107120	DN107120
c	38	32.6	17.0	651	1	AJ939538	AJ939538
	39	32.6	17.0	657	6	CP176758	CP176758
	40	32.6	17.0	662	6	CB478435	CB478435
c	41	32.6	17.0	683	6	CF787874	CF787874
c	42	32.6	17.0	695	6	CF175425	CF175425
c	43	32.6	17.0	706	9	BZ015309	BZ015309
c	44	32.6	17.0	822	2	B1185001	B1185001
c	45	32.6	17.0				

ALIGNMENTS

RESULT 1  
BP248044  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
COMMENT  
FEATURES  
source  
ORIGIN  
Query Match  
Best Local Similarity  
Matches  
QY  
DB

BP248044  
Sugano cdna library, embryo kidney Homo sapiens cDNA clone  
HKR09944, mRNA sequence.  
BP248044.1  
GI:52130323  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 582)  
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,  
Mizushima-Sugano, J., Nakai, K. and Sugano, S.  
Sequence comparison of human and mouse genes reveals a homologous  
block structure in the promoter regions  
Genome Res. 14 (9), 1711-1718 (2004)  
15342556  
Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp.  
Location/Qualifiers  
1..582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HKR09944"  
/tissue\_type="kidney"  
/cell\_line="293"  
/dev\_stage="embryo"  
/clone\_lib="Sugano cdna library, embryo kidney"

19.2%; Score 36.8; DB 3; Length 582;  
50.0%; Pred. No. 2;  
0; Mismatches 92; Indels 0; Gaps 0;  
Conservative  
1 AAAGCGGTATCTGTGACAGCGGCTGCAATACACTTGTGGATATGGAGAA 60  
151 AAGAACTGTAGCTTCTTCCTATTAGTGGATTACATGAGGATTGATAGATTAGGAAAA 210

```

Qy 61 AACAACTACTGCATAGGAATGCACATGGAGCACCAGGAGGAGTAATTACGGCTATTGC 120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
211 ACCATATATACAATTTAAAGAATGCAGATGGAAGGCCAGATAGGTGGTTGCCGAAGAGC 270
Qy 121 TACGAGATTGGGTGCTATTGCGAGGAGTTGTCCGATAGTACACCGACTTGGCCCTTTCT 180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
271 CTGGGAARACCAATTTAAACGAAATGATTCATCATAGTAGATATATTTCATGCGCCTTTT 330
Qy 181 AATA 184
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
331 CAAA 334

RESULT 2
BF640569 656 bp mRNA linear EST 19-DEC-2000
LOCUS NF031A09IN1F1068 Insect herbivory Medicago truncatula cDNA clone
DEFINITION NF031A09IN 5', mRNA sequence.
ACCESSION BF640569
VERSION BF640569
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 656)
AUTHORS Korth,K., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL Medicago truncatula insect herbivory library
COMMENT Unpublished (2000)
Contact: Korth K
Dept. of Plant Pathology
University of Arkansas
217 Plant Science Building, Fayetteville, AR 72701, USA
Tel: 501 575 5191
Fax: 501 575 7601
Email: kkorth@comp.uark.edu
Insert Length: 656 Std Error: 0.00
Plate: 031 Row: A Column: 09
Seq primer: TCACACGGAACACGCTATGAC.
FEATURES
source
location/Qualifiers
1..656
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF031A09IN"
/tissue_type="local and systemic leaves"
/dev_stages="mature"
/clone_lib="Insect herbivory"
/note="Vector: Lambda Zap; Library was produced from fully
expanded M. truncatula leaves of plants fed upon by
Spodoptera exigua (beet armyworm) for 24 hours. Systemic
(undamaged leaves from injured plants) and wounded leaves
were harvested and pooled."
ORIGIN
Query Match 18.3%; Score 35.2; DB 2; Length 656;
Best Local Similarity 52.0%; Pred. No. 6.5;
Matches 79; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 7 GGTATCTCGTGGACAACGCGGTGCAATATACACTTCTCGATGATTGGGAGAAAAA 66
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
464 GGTAAATTGGTGACACAGCTGGTGCAAAACCAATATCTTAGTTGGTATGGCTCAAC 523
Qy 67 TACTGCAATAGGGAATGCACATGGAAGCACCAGAGAGTAATTACGGCTATTGCTACGGA 126
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
524 TACATCTGTTCTTCTGCTCATGGCTTAATCAAAAGAGGAATGAAGGTTATTGCAGCTGG 583
Qy 127 TTGGGTGCTATTGCGAGGAGTTGTCGATAG 158
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 584 AATGAATCCTGTCCAAATTTGCTCGTGGGATTG 615
CNS003JOE 982 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR38N08 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL076634
AL076634.1 GI:4956111
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 982)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamooser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
location/Qualifiers
1..982
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR38N08"
/clone_lib="RPCI-98"
/note="end : TET3"
ORIGIN
Query Match 18.2%; Score 35; DB 10; Length 982;
Best Local Similarity 32.0%; Pred. No. 8.1;
Matches 49; Conservative 42; Mismatches 62; Indels 0; Gaps 0;

Qy 37 TACACTTCTGCGATATTGGGAGAAAAACAATATCTGCAATAGGGAATGCACATGGAAGCAC 96
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
714 KATWTWTTTGGWAAATTTGTGGCAAAAMWHAGGTGAAAACAGSTGCGANWGVVWAAA 655
Qy 97 CGAGGAGTAATTACGGCTATTGCTACGGAATTTGGGTGCTATTCCGAAGGATTTCCGAT 156
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
654 ARCGSGGAAAAAAAYGRCWGWSTCWTGCTTTCTYKTYTNGTVYMAAAAMWMAAMW 595
Qy 157 AGTACACCGACTTGGCCCTTTCTTAATAAAGA 189
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
594 TTVAAMAMMMWMTTTSAVACATWMAAAMAAACA 562

RESULT 4
AUI19061 820 bp mRNA linear EST 01-AUG-2002
LOCUS AUI19061 HEMBA1 Homo sapiens cDNA clone HEMBA1004961 5', mRNA
DEFINITION AUI19061 sequence.
ACCESSION AUI19061
VERSION AUI19061.1 GI:10934296
KEYWORDS EST.

```

**SOURCE**  
**ORGANISM**  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

**REFERENCE**  
**AUTHORS**  
 1 (bases 1 to 820)  
 Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,  
 Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and  
 Isogai, T.

**TITLE**  
**JOURNAL**  
**COMMENT**  
 HRI human cDNA project  
 Unpublished (2000)  
 Contact: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix  
 Research Institute; cDNA library construction: Department of  
 Virology, Institute of Medical Science, University of Tokyo, and  
 Helix Research Institute.

**FEATURES**  
**Source**  
 1..820  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="HEMBA1004961"  
 /tissue type="whole embryo, mainly head"  
 /dev stage="embryo, 10 weeks"  
 /clone\_lib="HEMBA1"  
 /note="Vector: pME18SFL3"

**ORIGIN**  
 Query Match 18.1%; Score 34.8; DB 1; Length 820;  
 Best Local Similarity 53.3%; Pred. No. 9;  
 Matches 72; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 42 TTGCTGGATATTGGGAGAAACAAATCTGCAATAGGGAATGCATGGAAGCACCAGG 101  
 |||||  
 Db 609 TTGCTGGATAGGATGACACAGATGGAGTANTAGCTTGCTGCTGGACTAATAGGG 668  
 |||||

QY 102 AGCTAATTACGCTATTGCTACGGATTGGCTCTATTGCGAGGATGTCGGATAGTAC 161  
 |||||

Db 669 ACTTATGTACAGCATTTGGGAAGGATGCTAGTAAATTAGTAGAAGCTAGCCCAATGATCC 728  
 |||||

QY 162 ACCGACTTGGCCCT 176  
 |||||

Db 729 CATGAATTGGTAAT 743  
 |||||

**RESULT 5**  
**DN839051**  
**LOCUS**  
**DEFINITION**  
 Smoc-1.05 A23.77 Selaginella moellendorffii cDNA library Smoc-1  
 Selaginella moellendorffii cDNA 5, mRNA sequence.

**ACCESSION**  
**VERSION**  
**KEYWORDS**  
**SOURCE**  
**ORGANISM**  
 DN839051 849 bp mRNA linear EST 09-JUL-2005  
 Smoc-1.05 A23.77 Selaginella moellendorffii cDNA library Smoc-1  
 Selaginella moellendorffii cDNA 5, mRNA sequence.

**REFERENCE**  
**AUTHORS**  
**TITLE**  
 Weng, J.-K., Tanurdic, M. and Chapple, C.  
 Functional analysis and comparative genomics of expressed sequence  
 tags from the lycophyte Selaginella moellendorffii

**JOURNAL**  
**PUBLISHED**  
**COMMENT**  
 BMC Genomics 6 (1), 85 (2005)  
 15938755  
 Contact: Chapple, C.  
 Department of Biochemistry  
 Purdue University

West Lafayette, IN 47907-1153, USA  
 Tel: 765 494 0494  
 Fax: 765 496 7213  
 Email: chapple@purdue.edu  
 PCR Primers  
 FORWARD: T7 Forward  
 BACKWARD: M13 Backward  
 Plate: Smoc-1.05 row: A column: 23  
 Seq primer: T7 Forward.

**FEATURES**  
**source**  
 1..849  
 Location/Qualifiers  
 /organism="Selaginella moellendorffii"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:88036"  
 /tissue type="whole plant"  
 /clone\_lib="Selaginella moellendorffii cDNA library  
 Smoc-1.05"  
 /note="Vector: pDR-LIB; Site 1: Sfil; Site 2: Sfil;  
 Library construction was performed according to  
 manufacture's (CLONTECH, Palo Alto, CA) recommended  
 protocol for Creator Smart cDNA Library Construction Kit."

**ORIGIN**  
 Query Match 18.1%; Score 34.8; DB 8; Length 849;  
 Best Local Similarity 57.3%; Pred. No. 9.1;  
 Matches 63; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 8 GTTATCTGTGACACAGCGGCTGCAATACACTTGTGGATATTGGGAGAAACAAAT 67  
 |||||

Db 376 GTAACTGGACGACCCAGGATCAGCATGTGCCAGGAGGCGATATGGAGCAGAAAG 435  
 |||||

QY 68 ACTGCAATAGGAATGCATCGAAGCACCAGGAGGTAATACGGCTAT 117  
 |||||

Db 436 AGTCTCTTTGAATATGCCCGGAACTTATGAGGTGGAATATACGGCTAT 485  
 |||||

**RESULT 6**  
**CF503342/c**  
**LOCUS**  
**DEFINITION**  
 MFI-0018U-V031-H07-U.B MFI-0018 Schistosoma mansoni cDNA clone  
 MFI-0018U-V031-H07.B similar to putative retrotransposon, mRNA  
 sequence.

**ACCESSION**  
**VERSION**  
**KEYWORDS**  
**SOURCE**  
**ORGANISM**  
 CF503342 340 bp mRNA linear EST 30-APR-2004  
 CF503342.1 GI:46896368  
 EST:  
 Schistosoma mansoni  
 Schistosoma mansoni  
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
 Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.

**REFERENCE**  
**AUTHORS**  
 DeMarco, R., Kowaltowski, A.T., Machado, A.A., Soares, M.B.,  
 Gargioni, C., Kawano, T., Rodrigues, V., Madeira, A.M.B.N.,  
 Wilson, R.A., Menck, C.F.M., Setubal, M.C., Dias-Neto, E., Leite, L.C.C.  
 and Verjovski-Almeida, S.  
 Sac1-1, -2 and -3 and Perere, four novel retrotransposons with high  
 transcriptional activities from the human parasite Schistosoma  
 mansoni

**TITLE**  
 J. Virol. 78 (6), 2967-2978 (2004)

**JOURNAL**  
**PUBLISHED**  
**COMMENT**  
 14990715  
 Contact: Dr. Sergio Verjovski-Almeida  
 Departamento de Bioquímica  
 Instituto de Química - Universidade de São Paulo  
 Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,  
 Brasil  
 Tel: +55-11-3091-2173  
 Fax: +55-11-3091-2186  
 Email: verjov@ig.usp.br  
 This sequence was derived from the FAPESP Schistosoma mansoni EST  
 Genome Project. All sequences in the project were assembled and  
 annotated. This entry and all the assembled sequences can be seen  
 in the following URL <http://bioinfo.ig.usp.br/schisto/>  
 Plate: MFI-0018U-V031 row: 7 column: H.

**FEATURES**  
 Location/Qualifiers

```

source
1. 340
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="MF1-0018U-V031-H07.B"
/sex="female"
/dev_stage="adult"
/lab_host="Mesocricetus auratus"
/clone_lib="MF1-0018"
/note="Vector: SureClone"

ORIGIN
Query Match      18.0%; Score 34.6; DB 6; Length 340;
Best Local Similarity 54.3%; Pred. No. 8.7;
Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 32 GCAATACACTTGGTGGATATTGGGAGAAAACAATACTGCAATAGGGAATGCACATGGA 91
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
177 GTAACCTCCACTTCTCGGATAAAATCTGAATTCAGCTATTGAATTAGCAATAGAAATCCGGA 118

Qy 92 AGCACCGAGGAGGTAAATACGGCTATTGCTACGGAATTTGGTGCTATTGCGAAGGATTGT 151
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117 GTCATAGAGGAATTAGACACATCAGTTCTTGAGGATATCGACATATTTTGGACTAATAAT 58

Qy 152 CCGATAGTA 160
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57 TCGATAGGA 49

RESULT 7
LOCUS BU552525
DEFINITION BU552525 918 bp mRNA linear EST 16-SEP-2002
5', mRNA sequence.
ACCESSION BU552525
VERSION BU552525.1 GI:22902797
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 918)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2774 row: k column: 06
High quality sequence stop: 736.
Location/Qualifiers
1. 918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6576318"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).

source
1. 1080
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 6 CAP GATE 1"
/note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Crown
(50%) and leaf (50%) tissues from wheat cultivar Norstar
after short exposure times to low temperature in the light
and in the dark. 12 mRNA populations were combined before
constructing the library. The first 6 populations: After 7
days of growth at 20C from wheat cultivar Norstar after
short exposure times to low temperature in the light and
in the dark. 12 mRNA populations were combined before
constructing the library. The first 6 populations: After 7
days of growth at 20, wheat plants were transferred to 4C
in the light. 1cm crown sections and green leaf tissue were
separately harvested after 1, 3, and 6 hours of low
temperature exposure. The last 6 populations: After 7 days
of growth at 20C, wheat plants were transferred to 4C in

```

```

ORIGIN
Note: this is a NIH_MGC Library."

Query Match      18.0%; Score 34.6; DB 5; Length 918;
Best Local Similarity 55.4%; Pred. No. 11;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 42 TTGCTGGATATTGGGAGAAACAATACTCAATAGGGAATGCACATGGAAGCACCGAGG 101
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
644 TTGCTGGAGATGAGGATCACACATGAGTAGTAGCTTTGTGCTGGACTAATAGGGG 703

Qy 102 AGGTAATTTACGGCTATTGCTACGATTTGGGTGCTATTGCCAAGGATTTGCCGATAGTAC 161
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
704 ACTTATGTACAGCATTTGGGAGGATGTACTTGAATTAGTAGAGCTAGGCCAATGATCC 763

Qy 162 A 162
Db |
764 A 764

RESULT 8
DR738811/c
LOCUS FGAS084028
DEFINITION FGAS084028 Triticum aestivum FGAS: Library 6 CAP GATE 1 Triticum
aestivum cDNA, mRNA sequence.
ACCESSION DR738811
VERSION DR738811.1 GI:70966654
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 1080)
AUTHORS Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
Penniket, C., Roach, J.L. and Sarhan, F.
Functional Genomics of Abiotic Stress in Wheat and Canola Crops
Unpublished (2003)
Contact: Patrick Gulick
Plant Molecular Biology
Concordia University, Department of Biology
7141 Sherbrooke St. West, Montreal, Quebec H4B 1R6, Canada
Tel: 514 848 2424 Ext 3407
Fax: 514 848 2881
Email: pgulick@alcor.concordia.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [105,489].
Plate: L6B202 row: N column: 17.
Location/Qualifiers
1..1080
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 6 CAP GATE 1"
/note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Crown
(50%) and leaf (50%) tissues from wheat cultivar Norstar
after short exposure times to low temperature in the light
and in the dark. 12 mRNA populations were combined before
constructing the library. The first 6 populations: After 7
days of growth at 20C from wheat cultivar Norstar after
short exposure times to low temperature in the light and
in the dark. 12 mRNA populations were combined before
constructing the library. The first 6 populations: After 7
days of growth at 20, wheat plants were transferred to 4C
in the light. 1cm crown sections and green leaf tissue were
separately harvested after 1, 3, and 6 hours of low
temperature exposure. The last 6 populations: After 7 days
of growth at 20C, wheat plants were transferred to 4C in

```

the dark. 1cm crown sections and green leaf tissue were separately harvested after 1, 3, and 6 hours of low temperature exposure. First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI. In addition, this library used a primer for second strand synthesis that annealed to an artificial sequence (RNA oligo) added before first strand synthesis. Therefore when sequences from EST generated from this library will be masked for vector and adaptor sequences, an additional masking step will have to be included to mask this RNA oligo that is common to all clones (sequence CAGCTGGAGCAGGACACTGCACATGCTGAGGACTAGAAA)."

## ORIGIN

Query Match 17.9%; Score 34.4; DB 8; Length 1080;  
Best Local Similarity 60.9%; Pred. No. 13;  
Matches 56; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 15 GGTGACAGACGGCTGCAATACACTTCTGCTATTTGGGAGAAACMAATACATGCAA 74

Db 359 GGGGTGCTATGGCGGTGTCATCTGCAGGGCGCTGTGTGTGAGGAGGAGAGACTGCAA 300

QY 75 TAGGCAATGCACATGCAACCGAGGAGTA 106

Db 299 CAGACACACATCTGGAGTCTCGAGGATATA 268

## RESULT 9

LOCUS CR899083/c  
DEFINITION Sus scrofa BBS, genomic survey sequence.  
ACCESSION CR899083  
VERSION CR899083.1 GI:56223580  
KEYWORDS GSS; Bac-end sequence BBS; Genome Survey Sequence.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 459)

AUTHORS Rogel-Gaillard, C., Bourgeaux, N., Billault, A., Vaiman, M. and Chardon, P.

TITLE Construction of a swine BAC library: application to the characterization and mapping of porcine type C endoviral elements

JOURNAL Cytogenet. Cell Genet. 85 (3-4), 205-211 (1999)

PUBMED 10449899

REFERENCE 2 (bases 1 to 459)

AUTHORS Chardon, P., Iannuccelli, N., Roig, A., Dossat, C., Demars, J., Rogel-Gaillard, C., Roy, A., Schibler, L. and Milan, D.

TITLE A physical map of the swine genome

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 459)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (18-NOV-2004) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

## FEATURES

source

1..459  
Location/Qualifiers  
/organism="Sus scrofa"  
/mol\_type="genomic DNA"  
/strain="Large White"  
/db\_xref="taxon:9823"  
/clone="BI0260D02"  
/sex="male"  
/cell\_type="fibroblast"  
/clone\_lib="SBAB"  
/notes="Genoscope sequence ID : IH0AAA26AA06FM1"

## ORIGIN

Query Match 17.8%; Score 34.2; DB 11; Length 459;

Best Local Similarity 62.1%; Pred. No. 12;

Matches 54; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 2 AAGACGGTTATCTGTGGACAGACGGCTGCAATACACTTGTGGATATTGGAGAAA 61

Db 455 AAGTGGGTAATGANGAACAACTTGACTGCAGACTCACTTCTATTATTCTGGAGAAA 396

QY 62 ACAATACTGCAATAGGAATGCACAT 88

Db 395 GTAGACATACTATGTTGTACGCATAT 369

## RESULT 10

LOCUS BQ864839

DEFINITION

QGC27N10.YG.ab1 QG\_ABCDI lettuce salinas Lactuca sativa cDNA clone

QGC27N10, mRNA sequence.

ACCESSION BQ864839

VERSION BQ864839.1

KEYWORDS

EST.

SOURCE

ORGANISM

Lactuca sativa

Lactuca sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; campanulids; Asterales; Asteraceae; Cichorioideae;

Cichorieae; Lactuca.

REFERENCE 1 (bases 1 to 490)

AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, W.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.

TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project

COMMENT http://compnoms.ucdavis.edu/

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Asmunsdon Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]

belongs to contig QG\_CA\_Contig3364, see http://cgdb.ucdavis.edu/

for details.

Plate: QGC27

row: N

column: 10.

Location/Qualifiers

1..490

/organism="Lactuca sativa"

/mol\_type="mRNA"

/cultivar="Salinas"

/db\_xref="taxon:4236"

/clone="QGC27N10"

/lab\_host="E. coli"

/clone\_lib="QG\_ABCDI lettuce salinas"

/note="Vector: pBRCNDSfiAB; The library was constructed

from 10 different sources of RNA from a single genotype.

Separate cDNAs were generated using primers that

incorporated unique 5' and 3' tags to distinguish each

source of RNA. cDNAs were then pooled, size-fractionated,

directionally cloned into a custom medium-copy vector and

transformations made with four size classes to minimize

size bias. Details of each source of RNA and library

construction can be obtained at http://cgdb.ucdavis.edu/

TAG\_SEQ=Not found"

## ORIGIN

Query Match 17.8%; Score 34.2; DB 5; Length 490;

Best Local Similarity 53.3%; Pred. No. 13;

Matches 72; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 3 AGACGGTTATCTGTGGACAGACGGCTGCAATACACTTGTGGATATTGGAGAAA 62

Db 37 AGCGGGTGGCAACACGAAACACACAGGTTCCACCACTGTGACGGCGTACGGCGGACGA 96

QY 63 CAAATACTGCAATAGGATGACATGGAAGCACCGAGGAGGTAAATTACGGCTATTGCTA 122

Db 97 CAACACTCCATATGGGACACACCAGGCGGTGTAGTGGCGGTGTGGGA 156  
 QY 123 CGGATTGGGTGCTA 137  
 Db 157 CGGGTTGGGTCCAA 171

RESULT 11  
 B0543170  
 LOCUS  
 DEFINITION BUS43170 893 bp mRNA linear EST 13-SEP-2002  
 AGENCOURT\_10338811 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6575177  
 5', mRNA sequence.

ACCESSION BUS43170  
 VERSION BUS43170.1 GI:22853653  
 KEYWORDS EST.  
 SOURCE

ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 893)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLCW2771 row: k column: 17  
 High quality sequence stop: 578.

FEATURES  
 source  
 1..893  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6575177"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 17.8%; Score 34.2; DB 5; Length 893;  
 Best Local Similarity 54.5%; Pred. No. 14;  
 Matches 66; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 42 TTGCTGGATATGGGAGAACCAATCTGCATAGGGAATGCACATGGACCGGAGG 101  
 Db 648 TTGCTGGAGATGAGGATCACACAGATGAGTAGTCTTGCTGCTGCACTAATAGGGG 707  
 QY 102 AGGTAATACGGCTATTGCTACGGAATTTGGGTGCTATTGCCAAGGATGTCGGATAGTAC 161  
 Db 708 ACTATGTACAGCATTTGNGAGGATGNTACTGATTAAGTAGAGCTAGGCCAATGATCC 767

QY 162 A 162  
 Db 768 A 768

RESULT 12

AY440472/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

AY440472 748 bp mRNA linear HTC 20-OCT-2004  
 Armigeres subalbatus ASAP ID: 39205 unknown mRNA sequence.

AY440472  
 AY440472.1 GI:42765501  
 HTC.

Armigeres subalbatus  
 Armigeres subalbatus

Armigeres subalbatus  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;  
 Culicidae; Culicinae; Culicini; Armigeres; Armigeres.

REFERENCE 1 (bases 1 to 748)  
 AUTHORS Bartholomay L.C., Cho W.-L., Rochelleau T.A., Boyle J.P., Beck E.T.,  
 Fuchs J.F., Liss P., Rusch M., Butler K.M., Wu R.C.-C., Lin S.-P.,  
 Kuo H.-Y., Tsao I.-Y., Huang C.-Y., Liu T.-T., Hsiao K.-J.,  
 Tsai S.-F., Yang U.-C., Nappi A.J., Perna N.T., Chen C.-C. and  
 Christensen B.M.

TITLE Description of the Transcriptomes of Immune Response-Activated  
 Hemocytes from the Mosquito Vectors Aedes aegypti and Armigeres  
 subalbatus  
 JOURNAL Infect. Immun. 72 (7), 4114-4126 (2004)  
 PUBMED 15213157

REFERENCE 2 (bases 1 to 748)

AUTHORS Bartholomay L.C., Cho W.-L., Rochelleau T.A., Boyle J.P., Beck E.T.,  
 Liss P., Rusch M., Fuchs J.F., Butler K.M., Wu R.C.-C., Kuo H.-K.,  
 Tsao I.-Y., Huang C.-Y., Hsiao K.-J., Tsai S.-F., Yang U.-C.,  
 Nappi A.J., Perna N.T., Chen C.-C. and Christensen B.M.

TITLE Direct Submission

JOURNAL Submitted (17-OCT-2003) Animal Health and Biomedical Sciences,  
 University of Wisconsin-Madison, 1656 Linden Dr., Madison, WI  
 53706, USA

COMMENT More information about this sequence is available in ASAP (A  
 Systematic Annotation Package for community analysis of genomes)  
 from the University of Wisconsin-Madison at  
<https://asap.ahabs.wisc.edu/annotation/php/logon.php>.

FEATURES  
 Location/Qualifiers

1..748  
 /organism="Armigeres subalbatus"  
 /mol\_type="mRNA"  
 /isolation\_source="perfused hemolymph of  
 bacteria-inoculated organisms at 1, 3, 6, 12, and 24  
 hours post-innoculation"  
 /db\_xref="taxon:124917"  
 /sex="female"  
 /cell\_type="hemocyte"  
 /tissue\_type="hemolymph"  
 /dev\_stage="adult"  
 /note="ASAP-UW Feature ID: 39204"  
 misc\_feature 1..748  
 /note="unknown; assembly of 12 EST sequences; ASAP-UW  
 Feature ID: 39205"  
 /db\_xref="ASAP:39205"

ORIGIN

Query Match 17.7%; Score 34; DB 4; Length 748;  
 Best Local Similarity 54.9%; Pred. No. 16;  
 Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 37 TACACTTGTCTGATTTGGGAGAAAACAAATCTCTGCAATAGGGAATGCACATGGAAAGCAC 96  
 Db 380 TACGATGCAGAGAAAGTGCCTAGAGTACTGTATCTGGGTTGCTGGGTTGTGGC 321  
 QY 97 CGAGGAGTAATACGGCTATTGCTACGGATTTGGGTGCTATTGCCAAGGATTTGCCAT 156  
 Db 320 TGATGAATTAATCGCGCTTTTGTGGCTGATGTGGGTGCTCGCGCAGCTGTGATCCGGT 261

QY 157 AG 158  
 Db 260 TG 259

RESULT 13  
 DR423779

LOCUS	DR423779	704 bp	mRNA	linear	EST 29-JUN-2005			
DEFINITION	nav30d12.y1 Human pterygium. Unnormalized (nav) Homo sapiens cDNA clone nav30d12 5', mRNA sequence.							
ACCESSION	DR423779							
VERSION	DR423779.1	GI:68325795						
KEYWORDS	EST.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE	1 (bases 1 to 704)							
AUTHORS	Wistow,G., Jaworski,C., Aryankalayil-John,M., Rowsey,J.J., Cox,C., Reid,T., Dushku,N. and Carper, D.							
TITLE	NRIBank analysis of Human pterygium							
JOURNAL	Unpublished (2005)							
COMMENT	Contact: Wistow G							
	Section on Molecular Structure and Function							
	National Eye Institute							
	6/331, NTH, Bethesda, MD 20892-2740, USA							
	Tel: 301 402 3452							
	Fax: 301 496 0078							
	Email: graeme@helix.nih.gov							
	Plate: 30 row: d column: 12							
	Seq primer: Universal M13 Reverse.							

FEATURES	SOURCE
1. The first part of the book is a history of the book of Daniel, from the time of the Babylonian captivity to the present day.	1. The first part of the book is a history of the book of Daniel, from the time of the Babylonian captivity to the present day.
2. The second part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	2. The second part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.
3. The third part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	3. The third part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.
4. The fourth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	4. The fourth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.
5. The fifth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	5. The fifth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.
6. The sixth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	6. The sixth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.
7. The seventh part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	7. The seventh part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.
8. The eighth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	8. The eighth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.
9. The ninth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	9. The ninth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.
10. The tenth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.	10. The tenth part of the book is a study of the book of Daniel, from the time of the Babylonian captivity to the present day.

```

Location/Qualifiers
1. .704
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="nav30d12"
/tissue_type="Pterygium"
/dev_stage="Adult"
/lab_host="EMDH108"
/clone_lib="Human pterygium. Unnormalized (nav)"
/notes="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted
from 9 pooled human pterygia. A directionally cloned cDNA
library in the pCMVSPORT6 vector (Invitrogen) was
constructed at Bioserve Biotechnology (Laurel MD)
essentially following the protocols of the Superscript
Plasmid System, full details of which are contained in the
manufacturer's Instruction manual
(http://www.lifetech.com/). First strand synthesis was
carried out using a Not I primer-adaptor
[5'-pGACATGTTTAGATCGGACGGCGGCCCT(7)15-3']. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed at
the NIH Intramural Sequencing Center (NISC). Analyzed data
available through http://neibank.nsl.nih.gov."
```

## ORIGIN

	Query Match	17.6%	Score 33.8;	DB 8;	Length 704;
	Best Local Similarity	52.5%;	Prod. No. 18;		
	Matches 74;	Conservative 0;	Mismatches 67;	Indels 0;	Gaps 0;
QY	46	TGGATATTGGGAGAAAAACAATACTGCAATAGGGAATGCATATGGAAGCACCGAGGAGGT	105		
DB	189	TAGAATTAGGAAAAAACCATATATCAATTAAAGATGCATATGGAAGGCCAGATAGGT	248		
QY	106	AATTACGGCTATTGCTACGGATTTGGGGTGCTATTGCGAAGGATTGTCCGATAGTACACCG	165		
DB	249	GGTTCCGAGNAGCGCTGGGAAACCAATTTAAACGAAATGATCTCTATCATAGTAGATAT	308		
QY	166	ACTTGGCCCTTTCTAAATAA	186		
DB	309	TTCATGGCCTTTTCAATCAA	329		

RESULT 14					
N75948/c					
LOCUS	N75948	377 bp	mRNA	linear	EST 29-MAR-1996
DEFINITION	za42c05.g1 Soares fetal liver spleen lNPLS Homo sapiens cDNA clone IMAGE:295208 3', similar to gb:K58957 TYROSINE-PROTEIN KINASE ATK (HUMAN); mRNA sequence.				

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

## REFERENCE AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 2.
2. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1.
3. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2.
4. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3.
5. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4.
6. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5.
7. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5, 6.
8. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5, 6, 7.
9. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5, 6, 7, 8.
10. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5, 6, 7, 8, 9.

## FEATURES

```

1. 377
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1240129"
/db_xref="taxon:9606"
/clone="IMAGE:295208"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_libs="S0ares fetal liver spleen INFUS"
/note="Organ: Liver and Spleen; Vector: p7T73D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACTGGAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."

```

## ORIGIN

Query Match	17.5%	Score 33.6;	DB 8;	Length 377;
Best Local Similarity	52.9%;	Pred. No. 18;		
Matches	72;	Conservative	0;	Mismatches 64; Indels 0; Gaps 0;
38	ACACTTCTGGGATATTGGGAGAAACAAATACTCAATAGGGAATGACATAGGAAGCACC	97		
207	AGACACAGACTGAATTTGGCATGAAATAATTTTTTAGGGGGAGGAATGAAATAGCCGCACA	148		
98	GAGGAGGTAAATPACGGCTATTGTACCGAATTTGGGTGCTATTGCCAAGGAATGTCGCGATA	157		
147	AGGGGTCCACAGACTCTTTGAGTAGGCAATTTGTTAGAGCTTTGGGGGGTGTGTGTGGGG	88		
158	GTACACCGACTTTGGGC	173		
87	GTGGACCGAAATTTGGC	72		

RESULT 15  
N57766/C

LOCUS	N57766	474 bp	mRNA	linear	EST 28-JAN-1997
DEFINITION	Y55603.a1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:246748 3', similar to gp:X58957 TYROSINE-PROTEIN KINASE ATK (HUMAN); mRNA sequence.				
ACCESSION	N57766				
VERSION	N57766.1	GI:1201656			



KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM  
REFERENCE  
AUTHORS  
1 (bases 1 to 474)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevaaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
and Marra, M.  
TITLE  
JOURNAL Generation and analysis of 280,000 human expressed sequence tags  
PUBMED Genome Res. 6 (9), 807-828 (1996)  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LInL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 372 Std Error: 0.00  
Seq primer: ml3 -40 forward  
High quality sequence stop: 342.  
FEATURES  
Location/Qualifiers  
1..474  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3795994"  
/db\_xref="taxon:9606"  
/clone="IMAGE:246748"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5', AACTGGAAGAATTAAATTAAGACTTTTCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN  
Query Match 17.5%; Score 33.6; DB 8; Length 474;  
Best Local Similarity 52.9%; Pred. No. 19;  
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 38 ACACCTTGCTGGATATTGGGAGAAACAATACTGCAATAGGATGCACATGGAGCACC 97  
DB 207 AGACAGACTGAATTTGCGATGAATAATTTTTTAGGAGGAGGATGTAATAGCGGCACA 148  
QY 98 GAGGAGGTAATTACGGCTATTGCTACGGATTGGGTGCTATTTCGAAGGATTGTCGGATA 157  
DB 147 AAGGGGTCCACAGCTCTTTGATGAGGCAATTTGGTAGAGCTTGGGGGGGTGTGTGTGGG 88  
QY 158 GTACACCGACTTGGCC 173  
DB 87 GTGGACCGAATTGGC 72

Search completed: January 22, 2006, 00:34:31  
Job time : 8948 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:23:21 ; Search time 303 Seconds

(without alignments)  
4223.172 Million cell updates/sec

Title: US-10-721-793-115

Perfect score: 192  
Sequence: 1 aaagacggttatcgttgga.....ccctttctaataaagatgc 192

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq 21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	192	14	ADY85771 DNA encod
2	192	100.0	323	14	ADY85769 DNA encod
3	190.4	99.2	192	14	ADY85775 DNA encod
4	190.4	99.2	323	14	ADY85773 DNA encod
5	156.8	81.7	192	14	ADY85807 DNA encod
6	156.8	81.7	320	14	ADY85805 DNA encod
7	155.2	80.8	192	14	ADY85851 DNA encod
8	155.2	80.8	320	14	ADY85849 DNA encod
9	152	79.2	192	14	ADY85703 DNA encod
10	152	79.2	258	14	ADY85701 DNA encod
11	150.4	78.3	192	14	ADY85699 DNA encod
12	150.4	78.3	254	14	ADY85697 DNA encod
13	145.4	75.7	198	14	ADY85747 DNA encod
14	145.4	75.7	323	14	ADY85745 DNA encod
15	126.8	66.0	189	14	ADY85767 DNA encod
16	126.8	66.0	311	14	ADY85765 DNA encod
17	124.6	64.9	323	14	ADY85757 DNA encod
18	124.6	64.9	323	14	ADY85761 DNA encod
19	123.2	64.2	192	14	ADY85763 DNA encod

20	123.2	64.2	192	14	ADY85759	ADY85759 DNA encod
21	123	64.1	195	14	ADY85803	ADY85803 DNA encod
22	123	64.1	198	14	ADY85831	ADY85831 DNA encod
23	123	64.1	323	14	ADY85801	ADY85801 DNA encod
24	123	64.1	323	14	ADY85829	ADY85829 DNA encod
25	119.8	62.4	198	14	ADY85723	ADY85723 DNA encod
26	119.8	62.4	198	14	ADY85727	ADY85727 DNA encod
27	119.8	62.4	198	14	ADY85815	ADY85815 DNA encod
28	119.8	62.4	198	14	ADY85819	ADY85819 DNA encod
29	119.8	62.4	322	14	ADY85721	ADY85721 DNA encod
30	119.8	62.4	322	14	ADY85725	ADY85725 DNA encod
31	119.8	62.4	323	14	ADY85817	ADY85817 DNA encod
32	119.8	62.4	323	14	ADY85813	ADY85813 DNA encod
33	118.2	61.6	195	14	ADY85795	ADY85795 DNA encod
34	118.2	61.6	198	14	ADY85811	ADY85811 DNA encod
35	118.2	61.6	319	14	ADY85809	ADY85809 DNA encod
36	118.2	61.6	323	14	ADY85793	ADY85793 DNA encod
37	116.6	60.7	195	14	ADY85739	ADY85739 DNA encod
38	116.6	60.7	274	14	ADY85737	ADY85737 DNA encod
39	116.6	60.7	323	14	ADY85753	ADY85753 DNA encod
40	115	59.9	195	14	ADY85827	ADY85827 DNA encod
41	115	59.9	195	14	ADY85823	ADY85823 DNA encod
42	115	59.9	323	14	ADY85821	ADY85821 DNA encod
43	115	59.9	323	14	ADY85825	ADY85825 DNA encod
44	113.4	59.1	195	14	ADY85799	ADY85799 DNA encod
45	113.4	59.1	195	14	ADY85679	ADY85679 DNA encod

ALIGNMENTS

RESULT 1

ADY85771

ID ADY85771 standard; DNA; 192 BP.

AC ADY85771;

DT 02-JUN-2005 (first entry)

DE DNA encoding the scorpion Cel3 toxin mature protein Seq 115.

KW gene; db; toxin; sodium channel; immunogenicity; antigen;

KW antibody production; venom; vaccine; diagnosis.

OS Centruroides elegans.

PN US2005065331-A1.

PD 24-MAR-2005.

PF 26-NOV-2003; 2003US-00721793.

PR 02-DEC-2002; 2002US-0430067P.

PA (UYME-) UNIV MEXICO NACIONAL AUTONOMA.

PI (SILA-) LAB SILANES SA DE CV.

PI Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;

PI Possani PLD;

DR WPI; 2005-252753/26.

DR P-PSDB; ADY85772.

PT Novel isolated scorpion toxin polypeptide, useful for preventing envenomation from scorpion stings, and as vaccine to prevent envenomation from venom of scorpions of genus Centruroides.

PS Claim 3; SEQ ID NO 115; 135pp; English.

XX This invention relates to novel scorpion toxin polynucleotides and the encoded proteins thereof having any one of 142 fully defined amino acid sequences given in the specification. Specifically, it refers to long chain toxins that block the sodium channels of excitable cells and also

CC short chain toxins that affect Erg type potassium channels. The present  
 CC invention describes immunogenic or antigenic compositions comprising a  
 CC scorpion toxin protein or fragment thereof, which can be used for the  
 CC generation of antibodies that are able to bind to and neutralize the in  
 CC vivo effects of scorpion venom. As such, they can be used in compositions  
 CC or appropriate recombinant fusion proteins in the development of vaccines  
 CC that can prevent envenomation from stings of scorpions of the genus  
 CC Centruroides. Furthermore, it provides a diagnostic method that uses an  
 CC immunogenic matrix for the purification of specific immunoglobulins  
 CC present in a sample that can determine the species of scorpion that has  
 CC stung an individual through the detection of identifying antibodies. In  
 CC addition, it provides methods that are useful for treating envenomation  
 CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
 CC mature protein of a sodium channel modifier toxin isolated from a  
 CC scorpion of the Centruroides genus, given in an exemplification of the  
 CC invention.

XX Sequence 192 BP; 59 A; 35 C; 53 G; 45 T; 0 U; 0 Other;

Query Match 100.0%; Score 192; DB 14; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 5e-55;  
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGCTGGACAGACGCGCTGCAATACACTTGTCTGGATATTGGGAGAA 60

DB 1 AAAGACGGTTATCTGCTGGACAGACGCGCTGCAATACACTTGTCTGGATATTGGGAGAA 60

QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGTC 120

DB 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGTC 120

QY 121 TACGGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180

DB 121 TACGGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180

QY 181 AATAAAGATGC 192

DB 181 AATAAAGATGC 192

# RESULT 2

ADY85769  
 ID ADY85769 standard; DNA; 323 BP.

XX AC ADY85769;

XX DT 02-JUN-2005 (first entry)

XX DE DNA encoding the full length scorpion Ce13 toxin protein Seq 113.

XX KW gene; ds; toxin; sodium channel; immunogenicity; antigen;  
 XX KW antibody production; venom; vaccine; diagnosis.

XX OS Centruroides elegans.

XX PN US2005065331-A1.

XX PD 24-MAR-2005.

XX PF 26-NOV-2003; 2003US-00721793.

XX PR 02-DEC-2002; 2002US-0430067P.

XX XX (UYME-) UNIV MEXICO NACIONAL AUTONOMA.  
 XX XX (SILA-) LAB SILANES SA DE CV.

XX PI Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
 XX Possani PLD;

XX DR WPI; 2005-252753/26.

XX DR P-PSDB; ADY85770.

XX PT Novel isolated scorpion toxin polypeptide, useful for preventing

PT envenomation from scorpion stings, and as vaccine to prevent envenomation  
 PT from venom of scorpions of genus Centruroides.

XX Claim 1; SEQ ID NO 113; 135pp; English.

XX This invention relates to novel scorpion toxin polynucleotides and the  
 CC encoded proteins thereof having any one of 142 fully defined amino acid  
 CC sequences given in the specification. Specifically, it refers to long  
 CC chain toxins that block the sodium channels of excitable cells and also  
 CC short chain toxins that affect Erg type potassium channels. The present  
 CC invention describes immunogenic or antigenic compositions comprising a  
 CC scorpion toxin protein or fragment thereof, which can be used for the  
 CC generation of antibodies that are able to bind to and neutralize the in  
 CC vivo effects of scorpion venom. As such, they can be used in compositions  
 CC or appropriate recombinant fusion proteins in the development of vaccines  
 CC that can prevent envenomation from stings of scorpions of the genus  
 CC Centruroides. Furthermore, it provides a diagnostic method that uses an  
 CC immunogenic matrix for the purification of specific immunoglobulins  
 CC present in a sample that can determine the species of scorpion that has  
 CC stung an individual through the detection of identifying antibodies. In  
 CC addition, it provides methods that are useful for treating envenomation  
 CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
 CC full length protein of a sodium channel modifier toxin isolated from a  
 CC scorpion of the Centruroides genus, given in an exemplification of the  
 CC invention.

XX Sequence 323 BP; 99 A; 56 C; 82 G; 86 T; 0 U; 0 Other;

Query Match 100.0%; Score 192; DB 14; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-55;  
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGCTGGACAGACGCGCTGCAATACACTTGTCTGGATATTGGGAGAA 60

DB 65 AAAGACGGTTATCTGCTGGACAGACGCGCTGCAATACACTTGTCTGGATATTGGGAGAA 124

QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGTC 120

DB 125 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGTC 184

QY 121 TACGGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180

DB 185 TACGGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 244

QY 181 AATAAAGATGC 192

DB 245 AATAAAGATGC 256

# RESULT 3

ADY85775

ID ADY85775 standard; DNA; 192 BP.

XX AC ADY85775;

XX DT 02-JUN-2005 (first entry)

XX DE DNA encoding the scorpion Ce13b toxin mature protein Seq 119.

XX KW gene; ds; toxin; sodium channel; immunogenicity; antigen;  
 XX KW antibody production; venom; vaccine; diagnosis.

XX OS Centruroides elegans.

XX PN US2005065331-A1.

XX PD 24-MAR-2005.

XX PF 26-NOV-2003; 2003US-00721793.

XX PR 02-DEC-2002; 2002US-0430067P.

XX XX (UYME-) UNIV MEXICO NACIONAL AUTONOMA.

PA (SILA-) LAB SILANES SA DE CV.  
 XX Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
 PI Possani PLD;  
 XX  
 XX WPI; 2005-252753/26.  
 DR P-PSDB; ADY85776.  
 DR  
 XX Novel isolated scorpion toxin polypeptide, useful for preventing  
 PT envenomation from scorpion stings, and as vaccine to prevent envenomation  
 PT from venom of scorpions of genus Centruroides.  
 XX  
 XX Claim 3; SEQ ID NO 119; 135pp; English.  
 XX  
 CC This invention relates to novel scorpion toxin polynucleotides and the  
 CC encoded proteins thereof having any one of 142 fully defined amino acid  
 CC sequences given in the specification. Specifically, it refers to long  
 CC chain toxins that block the sodium channels of excitable cells and also  
 CC short chain toxins that affect Erg type potassium channels. The present  
 CC invention describes immunogenic or antigenic compositions comprising a  
 CC scorpion toxin protein or fragment thereof, which can be used for the  
 CC generation of antibodies that are able to bind to and neutralize the in  
 CC vivo effects of scorpion venom. As such, they can be used in compositions  
 CC or appropriate recombinant fusion proteins in the development of vaccines  
 CC that can prevent envenomation from stings of scorpions of the genus  
 CC Centruroides. Furthermore, it provides a diagnostic method that uses an  
 CC immunogenic matrix for the purification of specific immunoglobulins  
 CC present in a sample that can determine the species of scorpion that has  
 CC stung an individual through the detection of identifying antibodies. In  
 CC addition, it provides methods that are useful for treating envenomation  
 CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
 CC mature protein of a sodium channel modifier toxin isolated from a  
 CC scorpion of the Centruroides genus, given in an exemplification of the  
 CC invention.  
 XX  
 XX Sequence 192 BP; 59 A; 36 C; 53 G; 44 T; 0 U; 0 Other;

Query Match 99.2%; Score 190.4; DB 14; Length 192;  
 Best Local Similarity 99.5%; Pred. No. 1.7e-54;  
 Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AAAGACGGTTATCTGTGGACAAAGACGGCTGCAAAATACACTTGTGGATATGGAGAA 60  
 DB 1 AAAGACGGTTATCTGTGGACAAAGACGGCTGCAAAATACACTTGTGGATATGGAGAA 60  
 QY 61 AACAAATACCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGTAAATACGGCTATTGC 120  
 DB 61 AACAAATACCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGTAAATACGGCTATTGC 120  
 QY 121 TAGCGATTTGGGTGCTATTGCGAAGGATTTGCGAGGATTTGCGAGTATGACCCGACTTGGCCCTTCT 180  
 DB 121 TAGCGATTTGGGTGCTATTGCGAAGGATTTGCGAGGATTTGCGAGTATGACCCGACTTGGCCCTTCT 180  
 QY 181 AATAAAGATGC 192  
 DB 181 AATAAAGATGC 192

RESULT 4  
 ADY85773  
 ID ADY85773 standard; DNA; 323 BP.  
 XX  
 XX ADY85773;  
 XX  
 XX 02-JUN-2005 (first entry)  
 XX  
 XX DNA encoding the full length scorpion Cel3b toxin protein Seq 117.  
 XX  
 XX gene; ds; toxin; sodium channel; immunogenicity; antigen;  
 XX antibody production; venom; vaccine; diagnosis.  
 XX  
 XX Centruroides elegans.  
 OS  
 XX

PN US2005065331-A1.  
 XX 24-MAR-2005.  
 XX  
 XX 26-NOV-2003; 2003US-00721793.  
 XX  
 XX 02-DEC-2002; 2002US-0430067P.  
 XX  
 XX (UYME-) UNIV MEXICO NACIONAL AUTONOMA.  
 PA (SILA-) LAB SILANES SA DE CV.  
 XX Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
 PI Possani PLD;  
 XX  
 XX WPI; 2005-252753/26.  
 DR P-PSDB; ADY85774.  
 DR  
 XX Novel isolated scorpion toxin polypeptide, useful for preventing  
 PT envenomation from scorpion stings, and as vaccine to prevent envenomation  
 PT from venom of scorpions of genus Centruroides.  
 XX  
 XX Claim 1; SEQ ID NO 117; 135pp; English.  
 XX  
 CC This invention relates to novel scorpion toxin polynucleotides and the  
 CC encoded proteins thereof having any one of 142 fully defined amino acid  
 CC sequences given in the specification. Specifically, it refers to long  
 CC chain toxins that block the sodium channels of excitable cells and also  
 CC short chain toxins that affect Erg type potassium channels. The present  
 CC invention describes immunogenic or antigenic compositions comprising a  
 CC scorpion toxin protein or fragment thereof, which can be used for the  
 CC generation of antibodies that are able to bind to and neutralize the in  
 CC vivo effects of scorpion venom. As such, they can be used in compositions  
 CC or appropriate recombinant fusion proteins in the development of vaccines  
 CC that can prevent envenomation from stings of scorpions of the genus  
 CC Centruroides. Furthermore, it provides a diagnostic method that uses an  
 CC immunogenic matrix for the purification of specific immunoglobulins  
 CC present in a sample that can determine the species of scorpion that has  
 CC stung an individual through the detection of identifying antibodies. In  
 CC addition, it provides methods that are useful for treating envenomation  
 CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
 CC full length protein of a sodium channel modifier toxin isolated from a  
 CC scorpion of the Centruroides genus, given in an exemplification of the  
 CC invention.  
 XX  
 XX Sequence 323 BP; 99 A; 57 C; 83 G; 84 T; 0 U; 0 Other;  
 QY  
 Query Match 99.2%; Score 190.4; DB 14; Length 323;  
 Best Local Similarity 99.5%; Pred. No. 2.1e-54;  
 Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AAAGACGGTTATCTGTGGACAAAGACGGCTGCAAAATACACTTGTGGATATGGAGAA 60  
 DB 65 AAAGACGGTTATCTGTGGACAAAGACGGCTGCAAAATACACTTGTGGATATGGAGAA 124  
 QY 61 AACAAATACCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGTAAATACGGCTATTGC 120  
 DB 125 AACAAATACCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGTAAATACGGCTATTGC 184  
 QY 121 TAGCGATTTGGGTGCTATTGCGAAGGATTTGCGAGGATTTGCGAGTATGACCCGACTTGGCCCTTCT 180  
 DB 185 TAGCGATTTGGGTGCTATTGCGAAGGATTTGCGAGGATTTGCGAGTATGACCCGACTTGGCCCTTCT 244  
 QY 181 AATAAAGATGC 192  
 DB 245 AATAAAGATGC 256  
 XX  
 XX  
 XX RESULT 5  
 XX ADY85807  
 XX ID ADY85807 standard; DNA; 192 BP.  
 XX  
 XX ADY85807;  
 XX

DT 02-JUN-2005 (first entry)  
XX DNA encoding the bark scorpion CseV1a toxin mature protein Seq 151.  
XX  
DE  
XX gene; ds; toxin; sodium channel; immunogenicity; antigen;  
KW antibody production; venom; vaccine; diagnosis.  
XX  
XX Centruroides sculpturatus.  
XX  
XX US2005065331-A1.  
XX  
XX PD 24-MAR-2005.  
XX  
XX 26-NOV-2003; 2003US-00721793.  
XX  
XX 02-DEC-2002; 2002US-0430067P.  
XX  
XX (UYME-) UNIV MEXICO NACIONAL AUTONOMA.  
XX (SILA-) LAB SILANES SA DE CV.  
XX  
XX Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
PI Possani PLD;  
PI  
XX WPI; 2005-252753/26.  
DR P-PSDB; ADY85808.  
DR  
XX Novel isolated scorpion toxin polypeptide, useful for preventing  
PT envenomation from scorpion stings, and as vaccine to prevent envenomation  
PT from venom of scorpions of genus Centruroides.  
XX  
XX Claim 3; SEQ ID NO 151; 135pp; English.  
XX  
XX This invention relates to novel scorpion toxin polynucleotides and the  
CC encoded proteins thereof having any one of 142 fully defined amino acid  
CC sequences given in the specification. Specifically, it refers to long  
CC chain toxins that block the sodium channels of excitable cells and also  
CC short chain toxins that affect Erg type potassium channels. The present  
CC invention describes immunogenic or antigenic compositions comprising a  
CC scorpion toxin protein or fragment thereof, which can be used for the  
CC generation of antibodies that are able to bind to and neutralize the in  
CC vivo effects of scorpion venom. As such, they can be used in compositions  
CC or appropriate recombinant fusion proteins in the development of vaccines  
CC that can prevent envenomation from stings of scorpions of the genus  
CC Centruroides. Furthermore, it provides a diagnostic method that uses an  
CC immunogenic matrix for the purification of specific immunoglobulins  
CC present in a sample that can determine the species of scorpion that has  
CC stung an individual through the detection of identifying antibodies. In  
CC addition, it provides methods that are useful for treating envenomation  
CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
CC mature protein of a sodium channel modifier toxin isolated from a  
CC scorpion of the Centruroides genus, given in an exemplification of the  
CC invention.  
XX  
SQ Sequence 192 BP; 62 A; 32 C; 51 G; 47 T; 0 U; 0 Other;  
Query Match 81.7%; Score 156.8; DB 14; Length 192;  
Best Local Similarity 88.5%; Pred. No. 4.9e-43;  
Matches 170; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 1 AAAGACGGTTATCTGTTGGACACAGCGGGCTGCAAAATACACTTGTCTGGATATTGGGAGAA 60  
DB 1 AAGACGGTTATCTAGTGGAAAGACGGGGCTGCAAAAGAGACTTGTCTACAAATTGGGAGAA 60  
QY 61 AACAAATCTGCAATAGGGAATGCATATGGAAGCAGCGAGGAGTAATTACGGCTATTGC 120  
DB 61 AACGATTTTTCATAGGGAATGCAATATGGAAGCAGACATAGGAGGTAGTTATGGCTATTTC 120  
QY 121 TACGGATTTGGTGTCTATTGGAGGATTTGTCGATAGTACACCGACTTGGCCCTTTCT 180  
DB 121 TACGATTTGGTGTCTATTGGAGGATTTGCCGATAGTACACAGACTTGGCCCTTCCT 180  
QY 181 AATAAAGATGC 192  
|||||

Db 181 AATAAAGATGC 192  
RESULT 6  
ADY85805  
ID ADY85805 standard; DNA; 320 BP.  
XX  
XX AC ADY85805;  
XX  
XX 02-JUN-2005 (first entry)  
XX  
XX DNA encoding the full length bark scorpion CseV1a toxin protein Seq 149.  
DE  
XX gene; ds; toxin; sodium channel; immunogenicity; antigen;  
KW antibody production; venom; vaccine; diagnosis.  
XX  
XX Centruroides sculpturatus.  
XX  
XX OS  
XX PN US2005065331-A1.  
XX  
XX PD 24-MAR-2005.  
XX  
XX 26-NOV-2003; 2003US-00721793.  
XX  
XX 02-DEC-2002; 2002US-0430067P.  
XX  
XX (UYME-) UNIV MEXICO NACIONAL AUTONOMA.  
XX (SILA-) LAB SILANES SA DE CV.  
XX  
XX Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
PI Possani PLD;  
PI  
XX WPI; 2005-252753/26.  
DR P-PSDB; ADY85806.  
DR  
XX Novel isolated scorpion toxin polypeptide, useful for preventing  
PT envenomation from scorpion stings, and as vaccine to prevent envenomation  
PT from venom of scorpions of genus Centruroides.  
XX  
XX Claim 1; SEQ ID NO 149; 135pp; English.  
XX  
XX This invention relates to novel scorpion toxin polynucleotides and the  
CC encoded proteins thereof having any one of 142 fully defined amino acid  
CC sequences given in the specification. Specifically, it refers to long  
CC chain toxins that block the sodium channels of excitable cells and also  
CC short chain toxins that affect Erg type potassium channels. The present  
CC invention describes immunogenic or antigenic compositions comprising a  
CC scorpion toxin protein or fragment thereof, which can be used for the  
CC generation of antibodies that are able to bind to and neutralize the in  
CC vivo effects of scorpion venom. As such, they can be used in compositions  
CC or appropriate recombinant fusion proteins in the development of vaccines  
CC that can prevent envenomation from stings of scorpions of the genus  
CC Centruroides. Furthermore, it provides a diagnostic method that uses an  
CC immunogenic matrix for the purification of specific immunoglobulins  
CC present in a sample that can determine the species of scorpion that has  
CC stung an individual through the detection of identifying antibodies. In  
CC addition, it provides methods that are useful for treating envenomation  
CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
CC full length protein of a sodium channel modifier toxin isolated from a  
CC scorpion of the Centruroides genus, given in an exemplification of the  
CC invention.  
XX  
SQ Sequence 320 BP; 100 A; 51 C; 80 G; 89 T; 0 U; 0 Other;  
Query Match 81.7%; Score 156.8; DB 14; Length 320;  
Best Local Similarity 88.5%; Pred. No. 5.9e-43;  
Matches 170; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 1 AAAGACGGTTATCTGTTGGACACAGCGGGCTGCAAAATACACTTGTCTGGATATTGGGAGAA 60  
DB 62 AAGACGGTTATCTAGTGGAAAGACGGGGCTGCAAAAGAGACTTGTCTACAAATTGGGAGAA 121  
QY 61 AACAAATCTGCAATAGGGAATGCATATGGAAGCAGCGAGGAGTAATTACGGCTATTGC 120

Db 122 AACGATTTTGCATAGGGAATGCAATGGAAGCACATAGGAGGTAGTTATGGCTATTTC 181  
QY 121 TACGATTTGGTGTCTATTGCGAAGATTGCGGATAGTACACCGACTTGGCCCTTTCT 180  
Db 182 TACGATTTGGTGTCTATTGCGAAGATTGCGGATAGTACACCGACTTGGCCCTTTCT 241  
QY 181 AATAAAAGATGC 192  
Db 242 AATAAAACATGC 253

RESULT 7  
ADY85851  
ID ADY85851 standard; DNA; 192 BP.  
XX AC ADY85851;  
XX DT 02-JUN-2005 (first entry)  
XX DE DNA encoding the bark scorpion CseI<sub>x</sub> toxin mature protein Seq 195.

XX KW gene; ds; toxin; sodium channel; immunogenicity; antigen;  
XX KW antibody production; venom; vaccine; diagnosis.  
XX OS Centruroides sculpturatus.

XX PN US2005065331-A1.  
XX PD 24-MAR-2005.  
XX PF 26-NOV-2003; 2003US-00721793.  
XX PR 02-DEC-2002; 2002US-0430067P.  
XX PA (UYME-) UNIV MEXICO NACIONAL AUTONOMA.  
XX PA (SILA-) LAB SILANES SA DE CV.  
XX PI Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
XX PI Possani PLD;  
XX WPI: 2005-252753/26.  
XX P-PSDB; ADY85852.

XX Novel isolated scorpion toxin polypeptide, useful for preventing  
PT envenomation from scorpion stings, and as vaccine to prevent envenomation  
PT from venom of scorpions of genus Centruroides.

XX Claim 3; SEQ ID NO 195; 135pp; English.

XX This invention relates to novel scorpion toxin polynucleotides and the  
CC encoded proteins thereof having any one of 142 fully defined amino acid  
CC sequences given in the specification. Specifically, it refers to long  
CC chain toxins that block the sodium channels of excitable cells and also  
CC short chain toxins that affect Erg type potassium channels. The present  
CC invention describes immunogenic or antigenic compositions comprising a  
CC scorpion toxin protein or fragment thereof, which can be used for the  
CC generation of antibodies that are able to bind to and neutralize the in  
CC vivo effects of scorpion venom. As such, they can be used in compositions  
CC or appropriate recombinant fusion proteins in the development of vaccines  
CC that can prevent envenomation from stings of scorpions of the genus  
CC Centruroides. Furthermore, it provides a diagnostic method that uses an  
CC immunogenic matrix for the purification of specific immunoglobulins  
CC present in a sample that can determine the species of scorpion that has  
CC stung an individual through the detection of identifying antibodies. In  
CC addition, it provides methods that are useful for treating envenomation  
CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
CC mature protein of a sodium channel modifier toxin isolated from a  
CC scorpion of the Centruroides genus, given in an exemplification of the  
CC invention.

XX Sequence 192 BP; 62 A; 29 C; 54 G; 47 T; 0 U; 0 Other;

XX SQ

Query Match 80.8%; Score 155.2; DB 14; Length 192;  
Best Local Similarity 88.0%; Pred. No. 1.7e-42;  
Matches 169; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
QY 1 AAAGACGGTTATCTGTTGGCAAGACGGGCTGCAAAATACACTTGTCTGATATTGGGAGAA 60  
Db 1 AAGGAAGTTATCTGTTGGACGTTAAGGGCTGCAAAAAAATTTGCTGGAATTTGGAGAT 60  
QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCAACCGAGAGGTAAATTACGGCTATTGC 120  
Db 61 AACGATTATTGCAATAGGGAATGTAATGGAAGCACATAGGAGGTAGTTACGGCTATTGC 120  
QY 121 TACGATTTGGTGTCTATTGCGAAGATTGCGGATAGTACACCGACTTGGCCCTTTCT 180  
Db 121 TACGATTTGGTGTCTATTGCGAAGATTGCGGATAGTACACCGACTTGGCCCTTTCT 180  
QY 181 AATAAAAGATGC 192  
Db 181 AATAAAACATGC 192

#### RESULT 8

ADY85849

ID ADY85849 standard; DNA; 320 BP.

XX AC ADY85849;

XX DT 02-JUN-2005 (first entry)

XX DE DNA encoding the full length bark scorpion CseI<sub>x</sub> toxin protein Seq 193.

XX KW gene; ds; toxin; sodium channel; immunogenicity; antigen;

XX KW antibody production; venom; vaccine; diagnosis.

XX OS Centruroides sculpturatus.

XX PN US2005065331-A1.

XX PD 24-MAR-2005.

XX PF 26-NOV-2003; 2003US-00721793.

XX PR 02-DEC-2002; 2002US-0430067P.

XX PA (UYME-) UNIV MEXICO NACIONAL AUTONOMA.

XX PA (SILA-) LAB SILANES SA DE CV.

XX PI Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;

XX PI Possani PLD;

XX WPI: 2005-252753/26.

XX P-PSDB; ADY85850.

XX Novel isolated scorpion toxin polypeptide, useful for preventing

PT envenomation from scorpion stings, and as vaccine to prevent envenomation

PT from venom of scorpions of genus Centruroides.

XX Claim 1; SEQ ID NO 193; 135pp; English.

XX This invention relates to novel scorpion toxin polynucleotides and the

CC encoded proteins thereof having any one of 142 fully defined amino acid

CC sequences given in the specification. Specifically, it refers to long

CC chain toxins that block the sodium channels of excitable cells and also

CC short chain toxins that affect Erg type potassium channels. The present

CC invention describes immunogenic or antigenic compositions comprising a

CC scorpion toxin protein or fragment thereof, which can be used for the

CC generation of antibodies that are able to bind to and neutralize the in

CC vivo effects of scorpion venom. As such, they can be used in compositions

CC or appropriate recombinant fusion proteins in the development of vaccines

CC that can prevent envenomation from stings of scorpions of the genus

CC Centruroides. Furthermore, it provides a diagnostic method that uses an

CC immunogenic matrix for the purification of specific immunoglobulins

CC present in a sample that can determine the species of scorpion that has

CC stung an individual through the detection of identifying antibodies. In

CC addition, it provides methods that are useful for treating envenomation

CC from scorpion stings. This polynucleotide is a DNA sequence encoding the

CC mature protein of a sodium channel modifier toxin isolated from a

CC scorpion of the Centruroides genus, given in an exemplification of the

CC invention.

XX Sequence 192 BP; 62 A; 29 C; 54 G; 47 T; 0 U; 0 Other;

XX SQ

CC stung an individual through the detection of identifying antibodies. In  
 CC addition, it provides methods that are useful for treating envenomation  
 CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
 CC full length protein of a sodium channel modifier toxin isolated from a  
 CC scorpion of the Centruroides genus, given in an exemplification of the  
 CC invention.

XX  
 SQ Sequence 320 BP; 99 A; 49 C; 85 G; 87 T; 0 U; 0 Other;

Query Match 80.8%; Score 155.2; DB 14; Length 320;  
 Best Local Similarity 88.0%; Pred. No. 2.1e-42;  
 Matches 169; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGCTGACAAAGACGGCTGCAAAATACACTTGTGGATATGGGAGAA 60  
 DB 62 AAGGAAGGTTATCTGCTGACCTAAGGGCTGCAAAAAAATGCTGGAAATTTGGAGAT 121  
 QY 61 AACAAATACCTGCAATAGGATGACATGGAAGCACCAGGAGGTAATTACGGCTATTGC 120  
 DB 122 AACGATTATGCAATAGGATGTAATGGAAGCACATAGGAGGTAGTACGGCTATTGC 181  
 QY 121 TACGGATTGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCCCTTCT 180  
 DB 182 TACGGATTGGTGCTATTGCGAAGGATTGCCGATAGTACACAGACTTGGCCCCCTTCT 241  
 QY 181 AATAAAGATGC 192  
 DB 242 AATAAAGATGC 253

#### RESULT 9

ADY85703

ID ADY85703 standard; DNA; 192 BP.

AC ADY85703;

DT 02-JUN-2005 (first entry)

XX DNA encoding the bark scorpion Cex10 toxin mature protein Seq 47.

DE gene; ds; toxin; sodium channel; immunogenicity; antigen;  
 KW antibody production; venom; vaccine; diagnosis.

XX Centruroides exilicauda.

XX US2005065331-A1.

XX 24-MAR-2005.

XX 26-NOV-2003; 2003US-00721793.

XX 02-DEC-2002; 2002US-0430067P.

XX (UYME-) UNIV MEXICO NACIONAL AUTONOMA.  
 XX (SILA-) LAB SILANES SA DE CV.

XX Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
 XX Possani PLD;

XX WPI; 2005-252753/26.

XX P-PSDB; ADY85704.

XX Novel isolated scorpion toxin polypeptide, useful for preventing  
 PT envenomation from scorpion stings, and as vaccine to prevent envenomation  
 PT from venom of scorpions of genus Centruroides.

XX Claim 3; SEQ ID NO 47; 135pp; English.

XX This invention relates to novel scorpion toxin polynucleotides and the  
 CC encoded proteins thereof having any one of 142 fully defined amino acid  
 CC sequences given in the specification. Specifically, it refers to long  
 CC chain toxins that block the sodium channels of excitable cells and also  
 CC short chain toxins that affect Erg type potassium channels. The present

CC invention describes immunogenic or antigenic compositions comprising a  
 CC scorpion toxin protein or fragment thereof, which can be used for the  
 CC generation of antibodies that are able to bind to and neutralize the in  
 CC vivo effects of scorpion venom. As such, they can be used in compositions  
 CC or appropriate recombinant fusion proteins in the development of vaccines  
 CC that can prevent envenomation from stings of scorpions of the genus  
 CC Centruroides. Furthermore, it provides a diagnostic method that uses an  
 CC immunogenic matrix for the purification of specific immunoglobulins  
 CC present in a sample that can determine the species of scorpion that has  
 CC stung an individual through the detection of identifying antibodies. In  
 CC addition, it provides methods that are useful for treating envenomation  
 CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
 CC mature protein of a sodium channel modifier toxin isolated from a  
 CC scorpion of the Centruroides genus, given in an exemplification of the  
 CC invention.

SQ Sequence 192 BP; 57 A; 34 C; 52 G; 49 T; 0 U; 0 Other;

Query Match 79.2%; Score 152; DB 14; Length 192;  
 Best Local Similarity 87.0%; Pred. No. 2.1e-41;  
 Matches 167; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGCTGACAAAGACGGCTGCAAAATACACTTGTGGATATGGGAGAA 60  
 DB 1 AAGACGGTTATCTGCTGAGGTACGGCTGCAAAAAGTCTTGCTATAAATTTGGAGAA 60  
 QY 61 AACAAATACCTGCAATAGGATGACATGGAAGCACCAGGAGGTAATTACGGCTATTGC 120  
 DB 61 AACAAATCTGCAATAGGATGCAAAATGAAGCACCAGGAGGTAGTACGGCTATTGC 120  
 QY 121 TACGGATTGGTGCTATTGCGAAGGATTGCCGATAGTACACCGACTTGGCCCCCTTCT 180  
 DB 121 TATTTTTTTGGTGCTATTGCGAAGGATTGCCGAAAAGTACACCGACTTGGCCCCCTTCT 180  
 QY 181 AATAAAGATGC 192  
 DB 181 AATAAATCATGC 192

#### RESULT 10

ADY85701

ID ADY85701 standard; DNA; 258 BP.

AC ADY85701;

XX 02-JUN-2005 (first entry)

XX DNA encoding the full length bark scorpion Cex10 toxin protein Seq 45.

DE gene; ds; toxin; sodium channel; immunogenicity; antigen;  
 KW antibody production; venom; vaccine; diagnosis.

XX Centruroides exilicauda.

XX US2005065331-A1.

XX 24-MAR-2005.

XX 26-NOV-2003; 2003US-00721793.

XX 02-DEC-2002; 2002US-0430067P.

XX (UYME-) UNIV MEXICO NACIONAL AUTONOMA.  
 XX (SILA-) LAB SILANES SA DE CV.

XX Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
 XX Possani PLD;

XX WPI; 2005-252753/26.

XX P-PSDB; ADY85702.

XX Novel isolated scorpion toxin polypeptide, useful for preventing  
 PT envenomation from scorpion stings, and as vaccine to prevent envenomation

PT from venom of scorpions of genus *Centruroides*.  
 XX Claim 1; SEQ ID NO 45; 135pp; English.  
 XX  
 CC This invention relates to novel scorpion toxin polynucleotides and the  
 CC encoded proteins thereof having any one of 142 fully defined amino acid  
 CC sequences given in the specification. Specifically, it refers to long  
 CC chain toxins that block the sodium channels of excitable cells and also  
 CC short chain toxins that affect Erg type potassium channels. The present  
 CC invention describes immunogenic or antigenic compositions comprising a  
 CC scorpion toxin protein or fragment thereof, which can be used for the  
 CC generation of antibodies that are able to bind to and neutralize the in  
 CC vivo effects of scorpion venom. As such, they can be used in compositions  
 CC or appropriate recombinant fusion proteins in the development of vaccines  
 CC that can prevent envenomation from stings of scorpions of the genus  
 CC *Centruroides*. Furthermore, it provides a diagnostic method that uses an  
 CC immunogenic matrix for the purification of specific immunoglobulins  
 CC present in a sample that can determine the species of scorpion that has  
 CC stung an individual through the detection of identifying antibodies. In  
 CC addition, it provides methods that are useful for treating envenomation  
 CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
 CC full length protein of a sodium channel modifier toxin isolated from a  
 CC scorpion of the *Centruroides* genus, given in an exemplification of the  
 CC invention.  
 XX  
 SQ Sequence 258 BP; 79 A; 48 C; 64 G; 67 T; 0 U; 0 Other;  
 Query Match 79.2%; Score 152; DB 14; Length 258;  
 Best Local Similarity 87.0%; Pred. No. 2.3e-41;  
 Matches 167; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
 QY 1 AAGACGGTTATCTGGTGACAGACGGGCTGCAATACACTTGTGGATATTGGAGAA 60  
 DB 1 AAGACGGTTATCTGGTGAGGTACGGGCTGCAAAAAGTCTGTATAAATTGGAGAA 60  
 QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGC 120  
 DB 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGC 120  
 QY 121 TACGGATTGGTGCTATTGGGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 180  
 DB 121 TATTTTGGTGCTATTGGGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 180  
 QY 181 AATAAAGATGC 192  
 DB 181 AATAAATCATGC 192  
 RESULT 11  
 ADY85699  
 ID ADY85699 standard; DNA; 192 BP.  
 XX  
 AC ADY85699;  
 XX  
 DT 02-JUN-2005 (first entry)  
 XX  
 DE DNA encoding the bark scorpion Cex9 toxin mature protein Seq 43.  
 XX  
 KW gene; ds; toxin; sodium channel; immunogenicity; antigen;  
 KW antibody production; venom; vaccine; diagnosis.  
 XX  
 OS *Centruroides exilicauda*.  
 XX  
 XX US2005065331-A1.  
 XX  
 XX 24-MAR-2005.  
 XX  
 XX 26-NOV-2003; 2003US-00721793.  
 XX  
 XX 02-DEC-2002; 2002US-0430067P.  
 XX  
 XX (UYME-) UNIV MEXICO NACIONAL AUTONOMA.  
 XX (SILA-) LAB SILANES SA DE CV.  
 PA

XX Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;  
 PI Possani PLD;  
 XX  
 XX WPI; 2005-252753/26.  
 DR P-PSDB; ADY85700.  
 DR  
 PT Novel isolated scorpion toxin polypeptide, useful for preventing  
 PT envenomation from scorpion stings, and as vaccine to prevent envenomation  
 PT from venom of scorpions of genus *Centruroides*.  
 XX  
 XX Claim 3; SEQ ID NO 43; 135pp; English.  
 XX  
 CC This invention relates to novel scorpion toxin polynucleotides and the  
 CC encoded proteins thereof having any one of 142 fully defined amino acid  
 CC sequences given in the specification. Specifically, it refers to long  
 CC chain toxins that block the sodium channels of excitable cells and also  
 CC short chain toxins that affect Erg type potassium channels. The present  
 CC invention describes immunogenic or antigenic compositions comprising a  
 CC scorpion toxin protein or fragment thereof, which can be used for the  
 CC generation of antibodies that are able to bind to and neutralize the in  
 CC vivo effects of scorpion venom. As such, they can be used in compositions  
 CC or appropriate recombinant fusion proteins in the development of vaccines  
 CC that can prevent envenomation from stings of scorpions of the genus  
 CC *Centruroides*. Furthermore, it provides a diagnostic method that uses an  
 CC immunogenic matrix for the purification of specific immunoglobulins  
 CC present in a sample that can determine the species of scorpion that has  
 CC stung an individual through the detection of identifying antibodies. In  
 CC addition, it provides methods that are useful for treating envenomation  
 CC from scorpion stings. This polynucleotide is a DNA sequence encoding the  
 CC full length protein of a sodium channel modifier toxin isolated from a  
 CC scorpion of the *Centruroides* genus, given in an exemplification of the  
 CC invention.  
 XX  
 SQ Sequence 192 BP; 57 A; 35 C; 52 G; 48 T; 0 U; 0 Other;  
 Query Match 78.3%; Score 150.4; DB 14; Length 192;  
 Best Local Similarity 86.5%; Pred. No. 7.3e-41;  
 Matches 166; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
 QY 1 AAGACGGTTATCTGGTGACAGACGGGCTGCAATACACTTGTGGATATTGGAGAA 60  
 DB 1 AAGACGGTTATCTGGTGAGGTACGGGCTGCAAAAAGTCTGTATAAATTGGAGAA 60  
 QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGC 120  
 DB 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGC 120  
 QY 121 TACGGATTGGTGCTATTGGGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 180  
 DB 121 TATTTTGGTGCTATTGGGAAGGATTGTCGATAGTACACCGACTTGGCCCTTCT 180  
 QY 181 AATAAAGATGC 192  
 DB 181 AATAAATCATGC 192  
 RESULT 12  
 ADY85697  
 ID ADY85697 standard; DNA; 254 BP.  
 XX  
 AC ADY85697;  
 XX  
 DT 02-JUN-2005 (first entry)  
 XX  
 DE DNA encoding the full length bark scorpion Cex9 toxin protein Seq 41.  
 XX  
 KW gene; ds; toxin; sodium channel; immunogenicity; antigen;  
 KW antibody production; venom; vaccine; diagnosis.  
 XX  
 OS *Centruroides exilicauda*.  
 XX  
 XX US2005065331-A1.  
 XX



```
XX PD 24-MAR-2005.
XX PF 26-NOV-2003; 2003US-00721793.
XX PR 02-DEC-2002; 2002US-0430067P.
XX PA (UYME-) UNIV MEXICO NACIONAL AUTONOMA.
XX PI (SILA-) LAB SILANES SA DE CV.
XX PI Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;
XX PI Possani PLD;
XX DR WPI; 2005-252753/26.
XX DR P-PSDB; ADY85698.
XX PT Novel isolated scorpion toxin polypeptide, useful for preventing
XX PT envenomation from scorpion stings, and as vaccine to prevent envenomation
XX PT from venom of scorpions of genus Centruroides.
XX PS Claim 1; SEQ ID NO 41; 135pp; English.
XX CC This invention relates to novel scorpion toxin polynucleotides and the
XX CC encoded proteins thereof having any one of 142 fully defined amino acid
XX CC sequences given in the specification. Specifically, it refers to long
XX CC chain toxins that block the sodium channels of excitable cells and also
XX CC short chain toxins that affect Erg type potassium channels. The present
XX CC invention describes immunogenic or antigenic compositions comprising a
XX CC scorpion toxin protein or fragment thereof, which can be used for the
XX CC generation of antibodies that are able to bind to and neutralize the in
XX CC vivo effects of scorpion venom. As such, they can be used in compositions
XX CC or appropriate recombinant fusion proteins in the development of vaccines
XX CC that can prevent envenomation from stings of scorpions of the genus
XX CC Centruroides. Furthermore, it provides a diagnostic method that uses an
XX CC immunogenic matrix for the purification of specific immunoglobulins
XX CC present in a sample that can determine the species of scorpion that has
XX CC stung an individual through the detection of identifying antibodies. In
XX CC addition, it provides methods that are useful for treating envenomation
XX CC from scorpion stings. This polynucleotide is a DNA sequence encoding the
XX CC full length protein of a sodium channel modifier toxin isolated from a
XX CC scorpion of the Centruroides genus, given in an exemplification of the
XX CC invention.
XX SQ Sequence 254 BP; 79 A; 48 C; 63 G; 64 T; 0 U; 0 Other;
Query Match 78.3%; Score 150.4; DB 14; Length 254;
Best Local Similarity 86.5%; Pred. No. 8.2e-41;
Matches 166; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 AAAGACGGTTATCTGGTGACACAGACGGGCTGCAATATACACTTGTCTGATATTGGGAGAA 60
DB 1 AAGGACGGTTATCTGGTGAGGTCACGGGCTGCAAAAAGTCTTGTATATAATTGGGAGAA 60
QY 61 AACAAATCTGCAATAGGAATGCACATGGAGCAGCAGGAGGTAATTACGGCTATTGC 120
DB 61 AACAAATCTGCAATAGGAATGCACAAATGAAGCAGCAGGAGGTAATTACGGCTATTGC 120
QY 121 TACGATTTGGGTGCTATTGGCAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 121 TATTTTTTGGGTGCTATTGGCAAGGATTGCGGAGATTTGGCGAAGTACACCGACTTGGCCCTTCT 180
QY 181 AATAAAGATGC 192
DB 181 AATAAATATGC 192
RESULT 13
ADY85747
ID ADY85747 standard; DNA; 198 BP.
XX AC
XX AC ADY85747;
XX DT 02-JUN-2005 (first entry)

XX DE DNA encoding the Mexican scorpion Cnl0b toxin mature protein Seq 91.
XX KW gene; ds; toxin; sodium channel; immunogenicity; antigen;
XX KW antibody production; venom; vaccine; diagnosis.
XX OS Centruroides noxius.
XX PN US2005065331-A1.
XX PD 24-MAR-2005.
XX PF 26-NOV-2003; 2003US-00721793.
XX PR 02-DEC-2002; 2002US-0430067P.
XX PA (UYME-) UNIV MEXICO NACIONAL AUTONOMA.
XX PI (SILA-) LAB SILANES SA DE CV.
XX PI Corona VM, Garcia RMC, Gurrola BG, Valdez CNA, Becerril LB;
XX PI Possani PLD;
XX DR WPI; 2005-252753/26.
XX DR P-PSDB; ADY85748.
XX PT Novel isolated scorpion toxin polypeptide, useful for preventing
XX PT envenomation from scorpion stings, and as vaccine to prevent envenomation
XX PT from venom of scorpions of Genus Centruroides.
XX PS Claim 3; SEQ ID NO 91; 135pp; English.
XX CC This invention relates to novel scorpion toxin polynucleotides and the
XX CC encoded proteins thereof having any one of 142 fully defined amino acid
XX CC sequences given in the specification. Specifically, it refers to long
XX CC chain toxins that block the sodium channels of excitable cells and also
XX CC short chain toxins that affect Erg type potassium channels. The present
XX CC invention describes immunogenic or antigenic compositions comprising a
XX CC scorpion toxin protein or fragment thereof, which can be used for the
XX CC generation of antibodies that are able to bind to and neutralize the in
XX CC vivo effects of scorpion venom. As such, they can be used in compositions
XX CC or appropriate recombinant fusion proteins in the development of vaccines
XX CC that can prevent envenomation from stings of scorpions of the genus
XX CC Centruroides. Furthermore, it provides a diagnostic method that uses an
XX CC immunogenic matrix for the purification of specific immunoglobulins
XX CC present in a sample that can determine the species of scorpion that has
XX CC stung an individual through the detection of identifying antibodies. In
XX CC addition, it provides methods that are useful for treating envenomation
XX CC from scorpion stings. This polynucleotide is a DNA sequence encoding the
XX CC full length protein of a sodium channel modifier toxin isolated from a
XX CC scorpion of the Centruroides genus, given in an exemplification of the
XX CC invention.
XX SQ Sequence 198 BP; 67 A; 35 C; 51 G; 45 T; 0 U; 0 Other;
Query Match 75.7%; Score 145.4; DB 14; Length 198;
Best Local Similarity 87.7%; Pred. No. 3.7e-39;
Matches 171; Conservative 0; Mismatches 21; Indels 3; Gaps 1;
QY 1 AAAGACGGTTATCTGGTGACAAAGCACAGGCTGTAATAACAACCTGCTGATATTGGGA 60
DB 1 AAGGACGGTTATCTGGTGACAAAGCACAGGCTGTAATAACAACCTGCTGATATTGGGA 60
QY 58 GAAAAACAATACTGCATAGGAATGCACATGGAGCAGCAGGAGGTAATTACGGCTAT 117
DB 61 GAAAAACAATACTGCATAGGAATGCACATGGAGCAGGAGGTAATTACGGCTAT 120
QY 118 TGCTACGGATTTGGGTGCTATTGGCAAGGATTGTCGATAGTACACCGACTTGGCCCTT 177
DB 121 TGCTACGGATTTGGGTGCTATTGGCAAGGATTGTCGATAGTACACCGACTTGGCCCTT 180
QY 178 TCTAATAAAGATGC 192
DB 181 CCTAATAAAGATGC 195
```



Query.Match 66.0%; Score 126.8; DB 14; Length 189;

```
Best Local Similarity 82.8%; Pred. No. 7.9e-33;
Matches 159; Conservative 0; Mismatches 27; Indels 6; Gaps 1;

Qy 1 AAAGACGGTTATCTGGTGGACAAGACGGGCTGCAAAATACACTTGTCTGGATATTGGGAGAA 60
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4 AAGGACGGTTATCTGGTGAACAAGACGGGCTGCAAAATACAATTGCTGGATATTGGGAGAA 63
Qy 61 AACAAATAGTGCNATAGGGAATGCAATGGAAGCACCGAGGAGGTAAATTACGGCTATTGC 120
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
64 AACAAATAGTGCNATTCGGAATGC-----AAAGAGGTAGGTGCTGGTTACGGCTATTGC 117
Qy 121 TACGGATTGGGTGCTATTGCGAAGGATTGTCCGATAGTACACCGACTTGGCCCTTTCT 180
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
118 TATGCTTTGGGTGCTATTGGAAGGATTACCCGAAGCGTACTGACCTGGCCCTTTCT 177
Qy 181 AATAAAGATGC 192
Db ||||||| |||
178 GATAAACATGC 189
```

Search completed: January 21, 2006, 21:36:21  
Job time : 303 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2006, 21:28:41 ; Search time 1730 Seconds  
(without alignments)

6308.637 Million cell updates/sec

Title: US-10-721-793-115

Perfect score: 192

Sequence: 1 aaagacggttatctgggtgga.....ccctttctaataaaagatgc 192

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_in:\*

3: gb\_env:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pr:\*

9: gb\_ro:\*

10: gb\_sts:\*

11: gb\_sy:\*

12: gb\_un:\*

13: gb\_vi:\*

14: gb\_hcg:\*

15: gb\_pi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158.4	82.5	192	2	AY351300 Centruroi
2	156.8	81.7	192	2	AY351297 Centruroi
3	156.8	81.7	319	2	CUDNATOXC
4	156.8	81.7	320	2	AF338460 Centruroi
5	155.2	80.8	192	2	AY351298 Centruroi
6	155.2	80.8	192	2	AY351299 Centruroi
7	155.2	80.8	320	2	AF338448 Centruroi
8	155.2	80.8	344	2	S81093 Centruroi
9	152	79.2	258	2	AY649868 Centruroi
10	150.4	78.3	254	2	AY649867 Centruroi
11	134.4	70.0	192	2	AY351302 Centruroi
12	134.4	70.0	192	2	AY351303 Centruroi
13	132.8	69.2	192	2	AY351301 Centruroi
14	131.2	68.3	192	2	AY351308 Centruroi
15	128	66.7	192	2	AY351307 Centruroi
16	124.6	64.9	335	2	CUDNATOXA
17	124.6	64.9	350	2	CUDNATOXB
18	123.2	64.2	192	2	AY351304 Centruroi

19	123.2	64.2	192	2	AY351306	Centruroi
20	123	64.1	323	2	AF338453	Centruroi
21	123	64.1	323	2	AF338462	Centruroi
22	121.6	63.3	192	2	AY351305	Centruroi
23	121.4	63.2	301	6	AR163001	Sequence
24	121.4	63.2	345	2	CNCN10	Y08270 C.noxius mr
25	121.4	63.2	345	6	AR163005	Sequence
26	119.8	62.4	320	2	AF338461	Centruroi
27	119.8	62.4	322	2	AF491129	Centruroi
28	119.8	62.4	322	2	AF491131	Centruroi
29	119.8	62.4	323	2	AF338456	Centruroi
30	119.8	62.4	323	2	AF338457	Centruroi
31	119.8	62.4	323	2	AF338458	Centruroi
32	118.2	61.6	316	2	AF491130	Centruroi
33	118.2	61.6	319	2	AF338459	Centruroi
34	118.2	61.6	323	2	AF338464	Centruroi
35	116.6	60.7	274	2	AF491134	Centruroi
36	115	59.9	323	2	AF338454	Centruroi
37	115	59.9	323	2	AF338455	Centruroi
38	113.4	59.1	261	2	AY649861	Centruroi
39	113.4	59.1	314	2	AY649862	Centruroi
40	113.4	59.1	323	2	AF338463	Centruroi
41	112.2	58.4	316	2	AF491133	Centruroi
42	112	58.3	261	2	AY649864	Centruroi
43	105.6	55.0	192	2	AY351310	Centruroi
44	104.8	54.6	273	2	AY649871	Centruroi
45	102.2	53.2	261	2	AY649863	Centruroi

#### ALIGNMENTS

RESULT 1  
AY351300  
LOCUS AY351300 192 bp DNA linear INV 31-DSC-2003  
DEFINITION Centruroides sculpturatus isolate CSEI beta-toxin gene, partial cds.

ACCESSION AY351300  
VERSION AY351300.1 GI:38017458

KEYWORDS Centruroides sculpturatus (bark scorpion)  
SOURCE Centruroides sculpturatus

ORGANISM Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthida; Buthoidea; Buthidae; Centruroides.

REFERENCE 1 (bases 1 to 192)

AUTHORS Zhu, S.

TITLE Alignment of beta-toxin nucleotide sequences

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 192)

AUTHORS Zhu, S.

TITLE Direct Submission

JOURNAL Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van Evenstraat, Leuven, Flanders 3000, Belgium

Location/Qualifiers

FEATURES

source

1..192

/organism="Centruroides sculpturatus"

/mol\_type="genomic DNA"

/isolate="CSEI"

/db\_xref="taxon:218467"

<1..>192

/product="beta-toxin"

<1..>192

/codon\_start=1

/product="beta-toxin"

/protein\_id="AA08035.1"

/db\_xref="GI:38017459"

/translation="KDGIVLTKTKCYKLGFNDFCNRECKWKHIGSGYCYGFG

CYCEGLPDSQTQWLPNKC"

ORIGIN

Query Match 82.5%; Score 158.4; DB 2; Length 192;

Best Local Similarity 89.1%; Pred. No. 6.1e-35;

Matches 171; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY	1	AAAGACGGTATCTGGTGCACAGACGGCTGCAATACACTTGGTATTTGGAGAA	60
Db	1	AAAGACGGTATCTAGTGGAAAGACGGCTGCAAAAAGACTTGTCAAAATTTGGAGAA	60
QY	61	AACAATACTGCAATAGGAATGCATGGAAGACCGAGGAGTAAATTACGGCTATTGC	120
Db	61	AACGATTTTGCATAGGAATGCAATGGAGACATAGGAGTAGTATTGGCTATTGC	120
QY	121	TACGATTGGTGTCTATTGGAGAGTATCCGATAGTACACCGACTTGGCCCTTTCT	180
Db	121	TACGATTGGTGTCTATTGGAGAGTATCCGATAGTACACCGACTTGGCCCTTTCT	180
QY	181	AATAAAGATGC 192	
Db	181	AATAAAGATGC 192	
RESULT 2			
AY351297			
LOCUS	AY351297	192 bp DNA linear	INV 31-DEC-2003
DEFINITION	Centruroides noxius isolate CnGTIV beta-toxin gene, partial cds.		
ACCESSION	AY351297		
VERSION	AY351297.1	GI:38017452	
KEYWORDS			
SOURCE	Centruroides noxius (Mexican scorpion)		
ORGANISM			
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;		
TITLE	Buthida; Buthoidea; Centruroides.		
JOURNAL	1 (bases 1 to 192)		
REFERENCE	Zhu,S		
AUTHORS	Alignment of beta-toxin nucleotide sequences		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 192)		
REFERENCE	Zhu,S.		
AUTHORS	Direct Submission		
TITLE	Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van		
JOURNAL	Evenstraat, Leuven, Flanders 3000, Belgium		
FEATURES	Location/Qualifiers		
source	1..192		
	/organism="Centruroides noxius"		
	/mol_type="genomic DNA"		
	/isolate="CnGTIV"		
	/db_xref="taxon:6878"		
	<1..>192		
	/product="beta-toxin"		
	<1..>192		
	/codon_start=1		
	/product="beta-toxin"		
	/protein_id="AAR08032.1"		
	/db_xref="GI:38017453"		
	/translation="KDGVLVDVKGCKKNCKYKLGENDYCNRECKMKHRRGSGYCYGFG		
	CYCEGLSDSTPTWPLPKRC"		
mRNA			
CDS			
ORIGIN			
Query Match	81.7%;	Score 156.8;	DB 2; Length 192;
Best Local Similarity	88.5%;	Pred. NO. 1.8e-34;	
Matches 170;	Conservative 0;	Mismatches 22;	Indels 0; Gaps 0;
QY	1	AAAGACGGTATCTGGTGCACAGACGGCTGCAATACACTTGGTATTTGGAGAA	60
Db	1	AAAGACGGTATCTAGTGGAAAGACGGCTGCAAAAAGACTTGTCAAAATTTGGAGAA	60
QY	61	AACAATACTGCAATAGGAATGCATGGAAGACCGAGGAGTAAATTACGGCTATTGC	120
Db	61	AACGATTTTGCATAGGAATGCAATGGAGACATAGGAGTAGTATTGGCTATTGC	120
QY	121	TACGATTGGTGTCTATTGGAGAGTATCCGATAGTACACCGACTTGGCCCTTTCT	180
Db	121	TACGATTGGTGTCTATTGGAGAGTATCCGATAGTACACCGACTTGGCCCTTTCT	180
QY	181	AATAAAGATGC 192	
Db	181	AATAAAGATGC 192	
QY	1	AAAGACGGTATCTGGTGCACAGACGGCTGCAATACACTTGGTATTTGGAGAA	60
Db	1	AAAGACGGTATCTAGTGGAAAGACGGCTGCAAAAAGACTTGTCAAAATTTGGAGAA	60
QY	61	AACAATACTGCAATAGGAATGCATGGAAGACCGAGGAGTAAATTACGGCTATTGC	120
Db	61	AACGATTTTGCATAGGAATGCAATGGAGACATAGGAGTAGTATTGGCTATTGC	120
QY	121	TACGATTGGTGTCTATTGGAGAGTATCCGATAGTACACCGACTTGGCCCTTTCT	180
Db	121	TACGATTGGTGTCTATTGGAGAGTATCCGATAGTACACCGACTTGGCCCTTTCT	180
QY	181	AATAAAGATGC 192	
Db	181	AATAAAGATGC 192	

```
mat_peptide 57..251
misc_feature /product="Na+-channel blocking toxin"
252..257
/standard_name="carboxyl-processing signal"
polyA_site 319
ORIGIN
Query Match 81.7%; Score 156.8; DB 2; Length 319;
Best Local Similarity 88.5%; Pred. No. 1.7e-34;
Matches 170; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
1 AAAGACGGTTATCTCGTGGACAAAGACGGGCTGCAAAATACACTTGTCTGGATATTGGAGAA 60
2 AAGACGGTTATCTCGTGGACAAAGACGGGCTGCAAAAGAAATTTGCTATAAATTGGAGAA 116
3 AACAAATCTGCAATAGGAATGCACATGGAAGACCGGAGGTAAATTACGGCTATTGC 120
4 AACGATTATTGCAATAGGAATGCACATGGAAGACCGGAGGTAGTTACGGCTATTGC 176
5 TACGGATTGGGTGCTATTGCGAAGGATTGCGGATAGTACACCGACTTGGCCCTTTCT 180
6 TACGGATTGGGTGCTATTGCGAAGGATTGCGGATAGTACACCGACTTGGCCCTTTCT 241
181 AATAAAGATGC 192
237 AATAAAGATGC 248
RESULT 4
AF338460 320 bp mRNA linear INV 26-OCT-2001
LOCUS Centruroides exilicauda sodium-channel modifier toxin precursor
DEFINITION CBE1a mRNA, complete cds.
ACCESSION AF338460
VERSION AF338460.1 GI:16444991
KEYWORDS Centruroides exilicauda
SOURCE Centruroides exilicauda
ORGANISM Centruroides exilicauda
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
Buthida; Buthoidea; Buthidae; Centruroides.
REFERENCE 1 (bases 1 to 320)
AUTHORS Corona,M., Valdez-Cruz,N.A., Merino,E., Zurita,M. and Possani,L.D.
TITLE Genes and peptides from the scorpion Centruroides sculpturatus
Ewing, that recognize Na(+) channels
JOURNAL Toxicon 39 (12), 1893-1898 (2001)
PUBMED 11600153
REFERENCE 2 (bases 1 to 320)
AUTHORS Corona,M. and Possani,L.D.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-2001) Bioestructura y Reconocimiento Molecular,
Instituto de Biociencia, Universidad Nacional Autonoma de
Mexico, Av. Universidad 2001, Cuernavaca, Morelos 62210, Mexico
FEATURES
source
1..320
Location/Qualifiers
/organism="Centruroides exilicauda"
/mol_type="mRNA"
/db_xref="taxon:6879"
/country="USA: Tucson, Arizona"
/notes="authority: Ewing"
5..265
/codon_start=1
/product="sodium-channel modifier toxin precursor CBE1a"
/protein_id="AAU23428.1"
/db_xref="GI:16444992"
/translation="MNSLLMITACLVLTIGTWAKDGYLVKRTGCKKTCYKLGNDPCN
RECKWKHIGSYGYFYFGYCYCEGLPDSQTWPLPNKTCGK"
5..61
sig_peptide
62..253
mat_peptide
251..253
misc_feature /product="sodium-channel modifier toxin CBE1a"
/notes="amidation site"
ORIGIN
Query Match 81.7%; Score 156.8; DB 2; Length 319;
Best Local Similarity 88.5%; Pred. No. 1.7e-34;
Matches 170; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
1 AAAGACGGTTATCTCGTGGACAAAGACGGGCTGCAAAATACACTTGTCTGGATATTGGAGAA 60
2 AAGACGGTTATCTCGTGGACAAAGACGGGCTGCAAAAGAAATTTGCTATAAATTGGAGAA 116
3 AACAAATCTGCAATAGGAATGCACATGGAAGACCGGAGGTAAATTACGGCTATTGC 120
4 AACGATTATTGCAATAGGAATGCACATGGAAGACCGGAGGTAGTTACGGCTATTGC 176
5 TACGGATTGGGTGCTATTGCGAAGGATTGCGGATAGTACACCGACTTGGCCCTTTCT 180
6 TACGGATTGGGTGCTATTGCGAAGGATTGCGGATAGTACACCGACTTGGCCCTTTCT 241
181 AATAAAGATGC 192
237 AATAAAGATGC 248
RESULT 5
AF351298 192 bp DNA linear INV 31-DEC-2003
LOCUS Centruroides noxius isolate Cni beta-toxin Gene, partial cds.
DEFINITION Centruroides noxius isolate Cni beta-toxin Gene, partial cds.
ACCESSION AF351298
VERSION AF351298.1 GI:38017454
KEYWORDS Centruroides noxius (Mexican scorpion)
SOURCE Centruroides noxius (Mexican scorpion)
ORGANISM Centruroides noxius (Mexican scorpion)
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
Buthida; Buthoidea; Buthidae; Centruroides.
REFERENCE 1 (bases 1 to 192)
AUTHORS Zhu,S.
TITLE Alignment of beta-toxin nucleotide sequences
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 192)
AUTHORS Zhu,S.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van
Evenstraat, Leuven, Flanders 3000, Belgium
FEATURES
source
1..192
Location/Qualifiers
/organism="Centruroides noxius"
/mol_type="genomic DNA"
/isolate="Cni"
/db_xref="taxon:6878"
<1..>192
/product="beta-toxin"
<1..>192
/codon_start=1
/product="beta-toxin"
/protein_id="PAB08033.1"
/db_xref="GI:38017455"
/translation="KDGVLVDAGKKKCYLKGKNDYCNRECMKRGSGSYCYCYGFG
CYCEGLSDSTPTWPLPNKTC"
ORIGIN
Query Match 80.8%; Score 155.2; DB 2; Length 192;
Best Local Similarity 88.0%; Pred. No. 5.1e-34;
Matches 169; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
1 AAAGACGGTTATCTCGTGGACAAAGACGGGCTGCAAAATACACTTGTCTGGATATTGGAGAA 60
2 AAGACGGTTATCTCGTGGACAAAGACGGGCTGCAAAAGAAATTTGCTATAAATTGGAGAA 60
61 AACAAATCTGCAATAGGAATGCACATGGAAGACCGGAGGTAAATTACGGCTATTGC 120
61 AACGATTATTGCAATAGGAATGCACATGGAAGACCGGAGGTAGTTACGGCTATTGC 120
121 TACGGATTGGGTGCTATTGCGAAGGATTGCGGATAGTACACCGACTTGGCCCTTTCT 180
121 TACGGATTGGGTGCTATTGCGAAGGATTGCGGATAGTACACCGACTTGGCCCTTTCT 180
121 TACGGATTGGGTGCTATTGCGAAGGATTGCGGATAGTACACCGACTTGGCCCTTTCT 180
```

Qy 181 AATAAAGATGC 192  
 Db 181 AATAAAGATGC 192

RESULT 6  
 AY351299 192 bp DNA linear INV 31-DEC-2003  
 LOCUS Centruroides sculpturatus isolate Cse1x beta-toxin gene, partial cds.  
 DEFINITION  
 AY351299 AY351299.1 GI:38017456  
 ACCESSION  
 VERSION  
 KEYWORDS Centruroides sculpturatus (bark scorpion)  
 SOURCE Centruroides sculpturatus  
 ORGANISM Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthida; Buthoidea; Buthidae; Centruroides.

REFERENCE  
 1 (bases 1 to 192)  
 Zhu, S.  
 AUTHORS Alignment of beta-toxin nucleotide sequences  
 TITLE Unpublished  
 JOURNAL  
 REFERENCE 2 (bases 1 to 192)  
 Zhu, S.  
 AUTHORS Direct Submission  
 TITLE Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van Evenstraat, Leuven, Flanders 3000, Belgium  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1..192  
 /organism="Centruroides sculpturatus"  
 /mol\_type="genomic DNA"  
 /isolate="Cse1x"  
 /db\_xref="taxon:218467"  
 <1..>192  
 /product="beta-toxin"  
 <1..>192  
 /codon\_start=1  
 /product="beta-toxin"  
 /protein\_id="AAR08034.1"  
 /db\_xref="GI:38017457"  
 /translation="KEGYLVDVKGCKKNCKWKLGDNDYCNRECKWKHIGSGYCYGFG CYCEGLPDSQTWPLPNKTC"

ORIGIN  
 Query Match 80.8%; Score 155.2; DB 2; Length 192;  
 Best Local Similarity 88.0%; Pred. No. 5.1e-34;  
 Matches 169; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1 AAGACGGTTATCTGGTGACAAAGACGGGCTGCAAAATACACTTGTGGATATTGGAGAA 60  
 Db 1 AAGGAGGTTATCTGGTGACCTAAGGGCTGCAAAAAAATTTCTGGAATTTGGAGAT 60

Qy 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGC 120  
 Db 61 AACGATTATGCAATAGGGAATGTAATGGAAGCACATAGGAGGTAGTTACGGCTATTGC 120

Qy 121 TACGATTGGGTGCTATTGGGAAGGATTGCCGATAGTACACCGACTTGGCCCTTTCT 180  
 Db 121 TACGATTGGGTGCTATTGGGAAGGATTGCCGATAGTACACAGACTTGGCCCTTCT 180

Qy 181 AATAAAGATGC 192  
 Db 181 AATAAAGATGC 192

RESULT 7  
 AF338448 320 bp mRNA linear INV 26-OCT-2001  
 LOCUS Centruroides exilicauda sodium-channel modifier toxin precursor  
 DEFINITION Cse1x mRNA, complete cds.  
 SOURCE Centruroides exilicauda  
 ORGANISM Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthida; Buthoidea; Buthidae; Centruroides.

REFERENCE  
 1 (bases 1 to 344)  
 Vazquez, A., Tapia, J. V., Eliason, W. K., Martin, B. M., Labretton, F., Delepierre, M., Possani, L. D. and Becerril, B.

SOURCE Centruroides exilicauda  
 ORGANISM Centruroides exilicauda  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthida; Buthoidea; Buthidae; Centruroides.

REFERENCE  
 1 (bases 1 to 320)  
 Corona, M., Valdez-Cruz, N. A., Mexino, E., Zurita, M. and Possani, L. D.  
 AUTHORS Genes and peptides from the scorpion Centruroides sculpturatus  
 TITLE Ewing, that recognize Na(+) channels  
 JOURNAL Toxicon 39 (12), 1893-1898 (2001)  
 PUBMED 11600153  
 REFERENCE 2 (bases 1 to 320)  
 Corona, M. and Possani, L. D.  
 AUTHORS Direct Submission  
 TITLE Submitted (22-JAN-2001) Bioestructura y Reconocimiento Molecular, Instituto de Biotecnologia, Universidad Nacional Autonoma de Mexico, Av. Universidad 2001, Cuernavaca, Morelos 62210, Mexico  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1..320  
 /organism="Centruroides exilicauda"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6879"  
 /country="USA: Tucson, Arizona"  
 /note="authority: Ewing"  
 5..265  
 /codon\_start=1  
 /product="sodium-channel modifier toxin precursor Cse1x"  
 /protein\_id="AAL23416.1"  
 /db\_xref="GI:16444968"  
 /translation="MNSLLMITACLVLVIGTVWAKGYLVVDVKGCKKNCKWKLGDNDYCN RECKWKHIGSGYCYGFGCYCEGLPDSQTWPLPNKTCGKK"

sig\_peptide 5..61  
 mat\_peptide 62..253  
 misc\_feature 251..253  
 /product="sodium-channel modifier toxin Cse1x"  
 /note="amidation site"

ORIGIN  
 Query Match 80.8%; Score 155.2; DB 2; Length 320;  
 Best Local Similarity 88.0%; Pred. No. 5e-34;  
 Matches 169; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1 AAGACGGTTATCTGGTGACAAAGACGGGCTGCAAAATACACTTGTGGATATTGGAGAA 60  
 Db 62 AAGGAGGTTATCTGGTGACGTAAGGGCTGCAAAAAAATTTCTGGAATTTGGAGAT 121

Qy 61 AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGC 120  
 Db 122 AACGATTATTCGAATAGGGAATGTAATGGAAGCACATAGGAGGTAGTTACGGCTATTGC 181

Qy 121 TACGATTGGGTGCTATTGGGAAGGATTGCCGATAGTACACCGACTTGGCCCTTTCT 180  
 Db 182 TACGATTGGGTGCTATTGGGAAGGATTGCCGATAGTACACAGACTTGGCCCTTCT 241

Qy 181 AATAAAGATGC 192  
 Db 242 AATAAAGATGC 253

RESULT 8  
 S81093 344 bp mRNA linear INV 24-MAR-2000  
 LOCUS Centruroides noxius Na+ channel-specific toxin 1 mRNA, complete cds.  
 DEFINITION S81093 GI:1336779  
 ACCESSION S81093  
 VERSION S81093.1  
 KEYWORDS Centruroides noxius (Mexican scorpion)  
 SOURCE Centruroides noxius  
 ORGANISM Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthida; Buthoidea; Buthidae; Centruroides.

REFERENCE  
 1 (bases 1 to 344)  
 Vazquez, A., Tapia, J. V., Eliason, W. K., Martin, B. M., Labretton, F., Delepierre, M., Possani, L. D. and Becerril, B.

**TITLE** Cloning and characterization of the cDNAs encoding Na<sup>+</sup> channel-specific toxins 1 and 2 of the scorpion *Centruroides noxius* Hoffmann  
**JOURNAL** Toxicon 33 (9), 1161-1170 (1995)  
**PUBMED** 8585086  
**REMARK** GenBank staff at the National Library of Medicine created this entry [NCBI gibbon 175934] from the original journal article.

**FEATURES**  
 source

1..344  
 /location/Qualifiers  
 1..344  
 /organism="Centruroides noxius"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6878"  
 /note="Hoffmann"

gene

1..344  
 /gene="Na<sup>+</sup> channel-specific toxin 1, Cn1"

CDS

9..269  
 /note="Na<sup>+</sup> channel-specific toxin 1, Cn1"  
 /note="conceptual translation differs from that in the paper: Cn1"  
 /codon\_start=1  
 /product="Na<sup>+</sup> channel-specific toxin 1"  
 /protein\_id="AAB36085.2"  
 /db\_xref="GI:7322068"  
 /translation="MNSLIMITACFLVLTGWAKDGYLVDAKCKKNCYKLGKNDYCN  
 RECRMKRGSGVGYCVGFCYCEGLSDSTPTWPLPNKTCGSK"

**ORIGIN**

Query Match 80.8%; Score 155.2; DB 2; Length 344;  
 Best Local Similarity 88.0%; Pred. No. 4.9e-34;  
 Matches 169; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGGTGACAAAGCGGCTGCAATACACCTGCTGGATATTGGGAGAA 60  
 |||||  
 Db 66 AAGGACGGTTATCTGGTGACCAAGGCTGCAAAAGAAATGCTATAAATTGGGAAA 125  
 |||||  
 QY 61 AACAAATCTGCAATAGGAATGCATGGAAGACCGGAGGAGTAATACCGCTATTGTC 120  
 |||||  
 Db 126 AACGATTATTCGATAGGAGATGCAGATGAACCGGAGGAGTAGTACCGCTATTGC 185  
 |||||

QY 121 TACGGATTTGGTGCTATTGCGAAGGATGTCGATAGTACACCGACTTGGCCCTTTCT 180  
 |||||  
 Db 186 TACGGATTTGGTGCTATTGTCGAAGGATGTCGATAGTACACCGACTTGGCCCTTCT 245  
 |||||  
 QY 181 AATAAAGATGC 192  
 |||||  
 Db 246 AATAAAGATGC 257  
 |||||

**RESULT 9**

AY649868 258 bp mRNA linear INV 18-AUG-2004  
**LOCUS** Centruroides exilicauda Cex10 neurotoxin precursor (Cex10) mRNA,  
 partial cds.

**ACCESSION**

**VERSION**

**KEYWORDS**

**SOURCE**

**ORGANISM**

Centruroides exilicauda  
 Centruroides exilicauda  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 Buthida; Buthoidea; Buthidae; Centruroides.

**REFERENCE**

**AUTHORS**

**TITLE**

Valdez-Cruz,N.A., Davila,S., Licea,A., Corona,M., Zamudio,F.Z.,  
 Garcia-Valdes,J., Boyer,L. and Possani,L.D.  
 Biochemical, genetic and physiological characterization of venom  
 components from two species of scorpions: Centruroides exilicauda  
 Wood and Centruroides sculpturatus Ewing

**JOURNAL**

**REFERENCE**

**AUTHORS**

**TITLE**

Submitted (10-JUN-2004) Molecular Medicine and Bioprocesses,  
 Institute of Biotechnology UNAM, Av. Universidad 2001, Cuernavaca,

**FEATURES**  
 source

1..258  
 /location/Qualifiers  
 1..258  
 /organism="Centruroides exilicauda"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6879"  
 <1..258  
 /gene="Cex10"  
 <1..204  
 /gene="Cex10"  
 /codon\_start=1  
 /product="Cex10 neurotoxin precursor"  
 /protein\_id="AAT98001.1"  
 /db\_xref="GI:51234252"  
 /translation="KDGVLVEVTGCKKCYKLGKNCFCNRECKMKHRRGSGYCYVFFG  
 CYCEGLAESTPTWPLPNKSCGK"

gene

<1..258  
 /gene="Cex10"

CDS

<1..204  
 /gene="Cex10"  
 /codon\_start=1  
 /product="Cex10 neurotoxin precursor"  
 /protein\_id="AAT98001.1"  
 /db\_xref="GI:51234252"  
 /translation="KDGVLVEVTGCKKCYKLGKNCFCNRECKMKHRRGSGYCYVFFG  
 CYCEGLAESTPTWPLPNKSCGK"

**ORIGIN**

Query Match 79.2%; Score 152; DB 2; Length 258;  
 Best Local Similarity 87.0%; Pred. No. 4.2e-33;  
 Matches 167; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 AAAGACGGTTATCTGGTGACAAAGCGGCTGCAATACACCTGCTGGATATTGGGAGAA 60  
 |||||  
 Db 1 AAGGACGGTTATCTGGTGAGGCTGACGGCTGCAAAAAGTCTTGTCTATAAATTGGGAAA 60  
 |||||  
 QY 61 AACAAATCTGCAATAGGAATGCATGGAAGACCGGAGGAGTAATACCGCTATTGTC 120  
 |||||  
 Db 61 AACAAATCTGCAATAGGAATGCATGGAAGACCGGAGGAGTAGTACCGCTATTGTC 120  
 |||||

QY 121 TACGGATTTGGTGCTATTGCGAAGGATGTCGATAGTACACCGACTTGGCCCTTTCT 180  
 |||||  
 Db 121 TATTTTGGTGCTATTGCGAAGGATGTCGCGAAGTACACCGACTTGGCCCTTCT 180  
 |||||  
 QY 181 AATAAAGATGC 192  
 |||||  
 Db 181 AATAAATCATGC 192  
 |||||

**RESULT 10**

AY649867 254 bp mRNA linear INV 18-AUG-2004  
**LOCUS** Centruroides exilicauda Cex9 neurotoxin precursor (Cex9) mRNA,  
 partial cds.

**ACCESSION**

**VERSION**

**KEYWORDS**

**SOURCE**

**ORGANISM**

Centruroides exilicauda  
 Centruroides exilicauda  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 Buthida; Buthoidea; Buthidae; Centruroides.

**REFERENCE**

**AUTHORS**

**TITLE**

Valdez-Cruz,N.A., Davila,S., Licea,A., Corona,M., Zamudio,F.Z.,  
 Garcia-Valdes,J., Boyer,L. and Possani,L.D.  
 Biochemical, genetic and physiological characterization of venom  
 components from two species of scorpions: Centruroides exilicauda  
 Wood and Centruroides sculpturatus Ewing

**JOURNAL**

**REFERENCE**

**AUTHORS**

**TITLE**

Submitted (10-JUN-2004) Molecular Medicine and Bioprocesses,  
 Institute of Biotechnology UNAM, Av. Universidad 2001, Cuernavaca,  
 Mor 62250, Mexico

**FEATURES**

source

1..254  
 /organism="Centruroides exilicauda"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6879"  
 <1..254  
 /gene="Cex9"  
 <1..204

gene

CDS





```
RESULT 13
AY351301      192 bp      DNA      linear      INV 31-DEC-2003
DEFINITION   Centruiroides sculpturatus isolate CSEv2a beta-toxin gene, partial
LOCUS        cds.
ACCESSION   AY351301
VERSION     AY351301.1 GI:38017460
KEYWORDS    Centruiroides sculpturatus (bark scorpion)
SOURCE      Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
ORGANISM    Buthida; Buthoidea; Buthidae; Centruiroides.
REFERENCE   1 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Alignment of beta-toxin nucleotide sequences
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van
Evenstraat, Leuven, Flanders 3000, Belgium
FEATURES     source
             1..192
             /organism="Centruiroides sculpturatus"
             /mol_type="genomic DNA"
             /isolate="CSEv2a"
             /db_xref="taxon:218467"
             <1..>192
             /product="beta-toxin"
             <1..>192
             /codon_start=1
             /product="beta-toxin"
             /protein_id="AAR08036.1"
             /db_xref="GI:38017461"
             /translation="KEGYLVNSTGCKYGLKLGNEGDCCKEAKNQGSYGVCYAF"
             CWCGLPBSTPTVPLPKSC"
ORIGIN
Query Match      69.2%; Score 132.8; DB 2; Length 192;
Best Local Similarity 80.7%; Pred. No. 1.5e-27;
Matches 155; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 1 AAGACGGTTATCTGCTGACAGACGGCTGCAAAATACACTTGTCTGGATATTGGGAGAA 60
DB 1 AAGAAGGTTATCTGTTAAACACGACGGCTGCAAAATACAGTTGCTCTGAAATTGGGAGAA 60
QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGACCGAGGAGGTAATTACGGCTATTGC 120
DB 61 AAGCAGGCTGCGATAGGAATGCAAGCGAAGAACCAAGGAGGTAGTTACGGCTATTGC 120
QY 121 TAGCGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 121 TAGCGTTTGCATGCTGTTGCGAAGGTTTCCCGAAGGTACACCGACTTATCTCTCTCT 180
QY 181 AATAAAGATGC 192
DB 181 AATAAATCATGC 192
RESULT 14
AY351308      192 bp      DNA      linear      INV 31-DEC-2003
LOCUS        Centruiroides noxius isolate Cn10 beta-toxin gene, partial cds.
DEFINITION   Centruiroides noxius (Mexican scorpion)
ACCESSION   AY351308
VERSION     AY351308.1 GI:38017474
KEYWORDS    Centruiroides noxius
SOURCE      Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
ORGANISM    Buthida; Buthoidea; Buthidae; Centruiroides.
REFERENCE   1 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Alignment of beta-toxin nucleotide sequences
ORIGIN
Query Match      69.2%; Score 132.8; DB 2; Length 192;
Best Local Similarity 80.7%; Pred. No. 1.5e-27;
Matches 155; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 1 AAGACGGTTATCTGCTGACAGACGGCTGCAAAATACACTTGTCTGGATATTGGGAGAA 60
DB 1 AAGAAGGTTATCTGTTAAACACGACGGCTGCAAAATACAGTTGCTCTGAAATTGGGAGAA 60
QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGACCGAGGAGGTAATTACGGCTATTGC 120
DB 61 AAGCAGGCTGCGATAGGAATGCAAGCGAAGAACCAAGGAGGTAGTTACGGCTATTGC 120
QY 121 TAGCGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 121 TAGCGTTTGCATGCTGTTGCGAAGGTTTCCCGAAGGTACACCGACTTATCTCTCTCT 180
QY 181 AATAAAGATGC 192
DB 181 AATAAATCATGC 192
```

```
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van
Evenstraat, Leuven, Flanders 3000, Belgium
FEATURES     source
             1..192
             /organism="Centruiroides noxius"
             /mol_type="genomic DNA"
             /isolate="Cn10"
             /db_xref="taxon:6978"
             <1..>192
             /product="beta-toxin"
             <1..>192
             /codon_start=1
             /product="beta-toxin"
             /protein_id="AAR08043.1"
             /db_xref="GI:38017475"
             /translation="KEGYLVNSTGCKYKYNCLILGENKNKDMCKAKNQGSYGVCYKLA
             CWCGLPBSTPTPIPGKTC"
ORIGIN
Query Match      68.3%; Score 131.2; DB 2; Length 192;
Best Local Similarity 80.2%; Pred. No. 4.4e-27;
Matches 154; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 1 AAGACGGTTATCTGCTGACAGACGGCTGCAAAATACACTTGTCTGGATATTGGGAGAA 60
DB 1 AAGGAAGGTTATCTGTTGAACACAGCTGCTGTAATACACTGCTTGATATTGGGAGAA 60
QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGACCGAGGAGGTAATTACGGCTATTGC 120
DB 61 AACAAAATCTGATATGGAATGCAAGCGAAGAACCAAGGAGGTAGTTACGGCTATTGC 120
QY 121 TAGCGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 121 TACAAGCTTGATGCTGCTGCGAAGGTTTCCCGAAGGTACACCGACTTATCCCATCTCT 180
QY 181 AATAAAGATGC 192
DB 181 GGTAAACATGC 192
RESULT 15
AY351307      192 bp      DNA      linear      INV 31-DEC-2003
LOCUS        Centruiroides sculpturatus isolate CSEv1d beta-toxin gene, partial
DEFINITION   cds.
ACCESSION   AY351307
VERSION     AY351307.1 GI:38017472
KEYWORDS    Centruiroides sculpturatus (bark scorpion)
SOURCE      Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
ORGANISM    Buthida; Buthoidea; Buthidae; Centruiroides.
REFERENCE   1 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Alignment of beta-toxin nucleotide sequences
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van
Evenstraat, Leuven, Flanders 3000, Belgium
FEATURES     source
             1..192
             /organism="Centruiroides sculpturatus"
             /mol_type="genomic DNA"
             /isolate="CSEv1d"
             /db_xref="taxon:218467"
             <1..>192
             /product="beta-toxin"
ORIGIN
Query Match      68.3%; Score 131.2; DB 2; Length 192;
Best Local Similarity 80.2%; Pred. No. 4.4e-27;
Matches 154; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 1 AAGACGGTTATCTGCTGACAGACGGCTGCAAAATACACTTGTCTGGATATTGGGAGAA 60
DB 1 AAGGAAGGTTATCTGTTGAACACAGCTGCTGTAATACACTGCTTGATATTGGGAGAA 60
QY 61 AACAAATCTGCAATAGGGAATGCACATGGAAGACCGAGGAGGTAATTACGGCTATTGC 120
DB 61 AACAAAATCTGATATGGAATGCAAGCGAAGAACCAAGGAGGTAGTTACGGCTATTGC 120
QY 121 TAGCGATTGGGTGCTATTGCGAAGGATTGTCGATAGTACACCGACTTGGCCCTTTCT 180
DB 121 TACAAGCTTGATGCTGCTGCGAAGGTTTCCCGAAGGTACACCGACTTATCCCATCTCT 180
QY 181 AATAAAGATGC 192
DB 181 GGTAAACATGC 192
```

```
CDS      <1..>192
          /codon_start=1
          /product="beta-toxin"
          /protein_id="AAR08042.1"
          /db_xref="GI:38017473"
          /translation="KEGYLVNSTGCKYDCFWLGKNEHCDLECKAKNOGSGSYGYCYAFA
          CWCEGLPESTPTYPELENKSC"

ORIGIN
Query Match      66.7%; Score 128; DB 2; Length 192;
Best Local Similarity 79.2%; Pred. No. 3.7e-26;
Matches 152; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy      1  AAAGACGGTTATCTGGTGAAGACGGGCTGCAAAATACACTTGTGATATTTGGAGAA 60
Db      1  AAAGAAAGTTATCTGGTAAACAGCACGGGCTGCAAAATACGATTTGTTGGTGGGAAA 60

Qy      61  AACAAATCTGCAATAGGGAATGCACATGGAAGCACCGAGGAGGTAATTACGGCTATTGC 120
Db      61  AACGNACACTGGGATTTGGATGCAAGCGAAGAACCAAGGAGGTAGTTACGGGTATTGC 120

Qy      121  TACGGATTTGGTGTCTATTGGGAAGGATTGTCGGATAGTACACCGACTTGGCCCTTTCT 180
Db      121  TACGCTTTGCGATGCTGTGGTGGGAGGTTTGGCCGAAAGTACACCGACTTATCCCTTCCT 180

Qy      181  AATAAAGATGC 192
Db      181  AATAAATCATGC 192
```

Search completed: January 21, 2006, 22:05:18  
Job time : 1731 secs

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New). Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

***This Page Blank (uspto)***